STRATEGY INTERVENTION TO ENHANCE READING COMPREHENSION OF 15-YEAR-OLD STUDENTS IN MEXICO

Alaidde Berenice Villanueva Aguilera

Doctor of Philosophy

University of York
Education

February 2014

Abstract

Results from the Programme for International Student Assessment (PISA) places Mexico at the lower end of a list of 32 countries in the evaluation of students' reading ability. Research suggests that students' reading comprehension can be improved by the direct teaching of reading strategies. Previous studies analysing the effectiveness of strategy training have demonstrated that comprehension can be enhanced by making students aware of their thinking while reading, and by teaching them to monitor their understanding. However, these studies focus only on comprehension of texts in English language and it is unclear whether these findings can be transferred to the Mexican context. It might be hypothesised that the direct teaching of reading strategies might also foster the comprehension of Spanish texts but few studies have attempted to test this hypothesis empirically.

This thesis reports on a quasi-randomised experimental intervention that was carried out in a secondary school located in a low-income working class neighbourhood in Mexico City. All students (N=177) in their last year of compulsory education were selected to participate in the project. All participants were randomly allocated to one of two groups: a) strategy group (direct teaching of reading strategies) and b) active control group. Students in the strategy group were taught cognitive and metacognitive strategies to test whether this would improve their reading comprehension. The intervention was carried out over a 6-month time period. The researcher taught both groups during the first two months and introduced the strategies. Two Spanish teachers subsequently practiced the strategies with the students for another two months. No further teaching of strategies was provided in the last two months.

The study employed both quantitative and qualitative data analyses. Improvements in comprehension were assessed using a comprehension test developed on the basis of material from PISA assessments. A self-reported instrument helped to examine differences in the reported use of reading strategies and findings from this instrument were corroborated with evidence from a 'think aloud' verbal report. Reading comprehension was assessed before the intervention and at two and six months. Strategy use was elicited before the intervention and at two months.

The results suggest that both groups improved their reading comprehension after the intervention but only the strategy group maintained their gains over time. Students in the strategy group more often reported to be aware of and use reading strategies than their peers in the control group. Both findings are statistically significant. However, a correlation test found no statistical significant correlation between gain scores of self-reported use of strategies and reading comprehension. This highlights the importance of employing multiple outcome measures to assess the effectiveness of reading comprehension instruction.

Contents

List of Tables

List of Figures

Acknowledgements

I am grateful to a number of people for their help and support throughout my studies. In particular, I am grateful to my supervisor, Dr Emma Marsden, for her time, expertise, constant support, encouragement and guidance, as well as my partner, Nils Gutacker, and my family for their constant support and unconditional love.

This research was possible thanks to the financial support provided by the Mexican Council for Science and Technology (CONACyT).

Author's declaration

I hereby declare that I am the sole author of this thesis and that this thesis has not been submitted for examination at this or any other institution for another award. I understand that my thesis may be made electronically available to the public.

1. Introduction

1.1. The low level of reading comprehension among Mexican students and its impact on economic growth

According to the International Standard Classification of Education, *"general education is defined as programmes designed to develop learners' general knowledge, skills and competencies, as well as literacy and numeracy skills"* (UNESCO, 2012, p. 14). There is little doubt about the benefits that education brings to individuals and society. For example, one of the main causes of high productivity is the quality of the human capital provided through education. The Organization for Economic Co-operation and Development (OECD) predicts that, for every additional year of education, the economic output of a country increases by 3-6% (OECD, 2006). Furthermore, a *"country able to attain literacy scores 1% higher than the international average will achieve levels of labour productivity and GDP per capita 2.5% and 1.5% higher"* (OECD, 2006, p. 155). Therefore, improvements in educational attainment and literacy can be considered an investment in productivity and, ultimately, economic growth[1].

Literacy plays an important role in individuals' lives, as it provides them with the tools to perform efficiently in society. Regardless of their occupation, literacy allows them to obtain information, reflect on it and, based on their knowledge, interact with other members of society. The influence of literacy on individuals' lives and society is so important that literacy levels are evaluated internationally to provide nations with comparable performance data. The OECD, for example, has developed evaluations to unify the criteria that allow assessment across countries. The results of these evaluations inform governments about their performance and provide a global perspective on the situation of the country. In 2000, the OECD launched the Programme for International Student Assessment (PISA). Its main objective is to evaluate *"how well young adults near the end of compulsory schooling are prepared to meet the challenges of today's knowledge societies"* (OECD, 2002, p. 14). PISA measures

[1] OECD's analyses show that increases in labour productivity might be responsible for the GDP growth between 1994 and 2004.

the performance of 15-year-old students in the areas of science, mathematics and reading. In most OECD countries, students of this age are about to enter the job market or pursue further studies. PISA evaluates the students' knowledge in these areas as well as their ability to reflect on and apply such knowledge. It is a long-term project, with assessments every three years[2]. These cycles allow for both cross-sectional analysis and longitudinal analysis of any changes in education system performance.

Reading is evaluated in PISA because it is considered an essential skill for life as well as a necessary skill to succeed in academic life. Since the early stages, students are taught to obtain information from texts and the amount and complexity of the materials increase considerably as their studies advance. PISA results for reading have revealed a problem with reading comprehension in Mexico (INEE, 2006). Mexican students perform worse than many other countries with regard to reading. This is not only reflected by the average reading performance (422 points compared to the OECD average of 501 points (OECD, 2002)) but also by the fact that the number of students with low reading proficiency is very high (47% in 2006).

The Mexican government has recognised the problem and placed reading comprehension on the national agenda for education (e.g. Martínez, 2013). It has also provided schools with resources (e.g. additional books) to promote reading. These efforts, however, have not been reflected in improvements in the national and international evaluations. The type and quality of the activities developed by the teachers to promote reading in schools might be one of the factors that have contributed to the poor results. Treviño et al. (2006) collected data on reading activities from 905 schools in Mexico to identify the classroom activities used by teachers to develop reading comprehension. Their results show that over 40% of the teachers regard the use of questionnaires or summaries as the most important activity to develop reading. Other activities included listening to somebody reading a text aloud or discussing the content of a book. However, the teaching of reading comprehension strategies was not given sufficient attention. More than 50% of the teachers reported a lack of knowledge regarding students' reading development. They argued that the current training was too theoretical and that it failed to focus on activities that work in practice. Moreover, 47% of the teachers were unaware of the criteria used in national and international evaluations to evaluate reading comprehension. The study did not report on the actual teaching of reading comprehension strategies in Mexican schools.

[2]The PISA study is coming to an end and the last evaluation will take place in 2015.

Their low comprehension might prevent Mexican students from pursuing higher education, and may cause a sustained shortage of skilled labour and impede economic growth in one of the largest developing countries in Latin America. The present study seeks to establish the effectiveness of teaching reading comprehensions strategies, which may help to address the problem of poor reading comprehension amongst Mexican students.

1.2. Rationale for the study

Research has shown that reading comprehension can be developed through instruction in reading strategies (Block, 2004; Duffy, 2002; Pressley, 2000). Several studies have been carried out to test the effectiveness of reading strategies (see Chapter 3). However, these studies were predominantly carried out in English-speaking countries and it is unclear whether the results can be generalised to the Mexican context. The present study has been designed to evaluate the effectiveness of the direct teaching of cognitive and metacognitive strategies in improving the reading comprehension of 15-year-old Mexican students. It is important to recognise that this problem may also need to be tackled at an earlier stage than age 15 (e.g. in the home, and at primary school) in order to develop the use and awareness of reading strategies. However, this study was carried out with 15-year old students for several reasons: a) students are considered to be sufficiently mature and possess a good prior knowledge of various general topics, b) the relative ease of access to a specific secondary school in Mexico, c) there is no comprehension instruction at this stage of education, and d) PISA is an important and highly policy relevant indicator of reading comprehension in 15-year-old students. To enhance ecological validity, some materials from the PISA assessment formed the basis for one of the outcome measures of the current study.

1.3. The study in the context of the international educational policy agenda

In 1990, the United Nations Educational, Scientific and Cultural Organization (UNESCO) launched a programme entitled Education for All (EFA), which aims to provide high-quality basic education for children, adolescents and adults by 2015. EFA defines quality as 'learners' cognitive development' accompanied by the promotion of values, attitudes of responsible citizenships and creative and emotional development

(UNESCO, 2004). The 2004 EFA annual report (UNESCO, 2004) suggests focussing efforts on eight areas when developing policies to foster the quality of education:

1. *Teachers* - more and better trained
2. *Learning time* - on average, 850 to 1,000 hours of instruction per year
3. *Core subjects* - it emphasises the acquisition of literacy as a critical tool for mastering other subjects and reading as a priority area
4. *Pedagogy* - the use of new techniques plays an active role in students' learning
5. *Language* - instruction in students' first language
6. *Learning materials* - quality and quantity of text books
7. *Facilities* - classrooms and schools with adequate conditions
8. *Leadership* - school autonomy to take decisions based on its specific characteristics and needs

Points 1 and 3-5 were taken into consideration during the design of the present study. The intervention not only aimed to influence the students' awareness and use of reading comprehension strategies but also to train schoolteachers how to teach them.

1.4. Structure of this work

Chapter 2 of this work provides a contextualised rationale for the study. It includes a description of the educational system in Mexico, and a more detailed explanation of the PISA evaluation and its characteristics as well as a review of the results of other national and international evaluations. This is followed by a review of the results for Mexico and a comparison with other Latin-American countries. The current policy of the Mexican government is also described, which aims to tackle the comprehension problem. The importance of teacher education is highlighted as the study included the training of two teachers. Chapter 3 will review the literature and relevant theories, describe approaches to strategies and strategy instruction, and present studies conducted in the area of reading comprehension research. Chapter 4 will provide the methodology of the project, describing the sample, the materials used for instruction, and the outcome measurements (a comprehension test, a

word level reading test, a verbal report, and a self-reported strategy questionnaire). Chapter 5 reports the results of the analysis of one of the outcome measures, the reading comprehension test, followed by an analysis which explores the possibility that the students' comprehension might be poor due to their low level of reading skills. Chapter 6 reports the results of the analysis of a self-report instrument that investigated the students' perceptions of their use of reading strategies and the analysis of verbal report investigating the students' use of strategies in action. It also provides the results of a further analysis where the results of the three main outcome instruments were analysed together to achieve a more in depth understanding of the results and the factors that may explain them. Chapter 7 provides a further discussion of the main results. Chapter 8 provides a summary of the main findings, the limitations of the study, and implications for future practice and research.

2. The Mexican educational system and how its performance compares internationally

2.1. Education in Mexico[1]

The educational system in Mexico is divided into three stages: basic, middle and higher education[2]. Basic education includes kindergarten (4-6 year-olds), primary school (6-12 year-olds) and secondary school (12-15 year-old)[3]. Middle education is represented by general 'bachilleratos' (equivalent to GCSE) or technical education. Higher education includes upper technical education, bachelor degrees, and post-graduate studies. Basic education is regulated by the Secretariat of Public Education (SEP). This ministry is in charge of the design, update and delivery of basic education (hereafter referred to as 'public education'). The private sector is another provider of basic education. The difference between the two sectors lies in the fact that private education complements the syllabi provided by the SEP. Students in private education have extended schedules. During their extra hours, the students participate in extracurricular activities such as studying foreign languages, IT, and art. About 87.7% of students receive education in the public sector, whereas the remaining 12.3% attend private education (INEE, 2006a).

2.1.1. Type of schools

Mexico is the second most populous country in Latin America (with a population of approximately 118m, second only to Brazil, with nearly 200m inhabitants) and the fifth largest in size (about 1.9m km). Its diverse geographical characteristics result

[1] The numbers and percentages employed in this section are mainly taken from the INEEs annual reports. It is worth mentioning that, before 2002, there was no institution responsible for the assessment of education in Mexico. Therefore some of the data might be inaccurate or missing.

[2] See Appendix A.1 for the equivalents to the English system.

[3] These can vary e.g. if students do not progress to the next year.

in high levels of dispersion among its population[4] and an educational system with complex characteristics. Thus, different types of school can be identified across the country. The characteristics of each type of school depend mainly on its location; that is, whether the school is located within an urban or a rural area. Within primary education, there are general schools, community centres and schools for ethnic groups (indigenous people). Secondary education is offered in general schools, technical schools, schools for workers, 'telesecundarias' (education delivered through TV programmes) and community schools. The schools delivering private education form another type of school.

In the rural areas, special schools operate, which are called multigrade (multigrado) schools. This type of school offers educational services in areas where, due to the low number of students, it is impossible to provide one teacher per class. Multi-grade schools are run by two or three teachers who work with students from two or three different educational levels at the same time. In addition, the schools in these areas are usually poorly-equipped and the surrounding community suffers from poverty and marginalisation. About 44% of primary schools operate under such precarious conditions (INEE, 2006b).

2.1.2. School life expectancy

The National Institute for the Assessment of Education (INEE - Instituto Nacional para la Evaluación de la Educación) is in charge of: a) developing and facilitating national and international tests (e.g. PISA), b) assessing schools resources and processes, and c) evaluating the quality of education. Since 2002, the INEE has informed Mexican society (i.e. policy makers, the Secretariat of Education, principals, teachers, parents, etc.) about the quality of education provided in the public and private sectors. Quality is assessed based on the students performance in national and international evaluations and on economic and socio-cultural elements.

Although there has been a 25% decrease in the birth rate over the last 10 years, there are still 26 million children who should legally attend school (INEE, 2006b). At the age of six, about 91% of students are enrolled in primary education; however, this figure decreases to 63% in 6th grade (12-year-old students) (INEE, 2006b). In the case of secondary education, 77% of students attend this level. Unfortunately, the dropout rate among this group is approximately 78%. This rate might be the

[4]There are a great number of small communities with less than 100 inhabitants. Small communities are classified into four types: close to motorways, close to rural communities, isolated communities and close to cities.

result of insufficient geographic access to schools, since half of the secondary schools in the country (52%) are located in eight of the 31 states. An important additional challenge for any educational programme, particularly those related to literacy, is that Mexico is a highly multicultural and multilingual country with more than 60 ethnic groups,[5] with special schools providing opportunities to indigenous people.

It is known that high quality education is associated with school life expectancy rates (OECD, 2006). School life expectancy is defined as the total number of years of schooling that a child of a certain age can expect to attain in the future. Despite the governments efforts to improve school life expectancy (INEE, 2006a), there remains scope for improvement. With 13.4 years, Mexico ranks 29th out of 30 countries that submitted data to the OECD in a recent comparison of education expectancy (OECD, 2006). This is well below other Latin American countries such as Brazil (16.7 years) or Chile (15.0 years) as well as below the OECD average of 16.1 years.

Ranking	Country	School life expectancy
1	Denmark	18.2
2	Finland	18.1
3	Iceland	17.5
4	Germany	17.3
5	Portugal	17.1
11	Brazil	16.7
16	Spain	16.3
22	United Kingdom	15.3
24	United States	15.0
25	Chile	15.0
29	Mexico	13.4
	OECD average	16.1

Table 2.1.: Average school life expectancy in selected OECD countries and partner countries.
(Source: OECD (2006)) [6]

[5] An indigenous home is defined as one where the father, mother or another member of the household speaks a native language.

[6] Only OECD countries ranked above Mexico are presented.

2.2. National Evaluations

2.2.1. EXCALE

The INEE has developed the Quality and Educational Achievement Tests (EXCALE - Examenes de la Calidad y el Logro Educativo), which are designed to collect comparable data between cohorts, which serve to measure the quality of education in Mexico over time and identify trends in performance. The first cycle of EXCALE started in 2005 (with evaluations of Spanish and Mathematics) and has since been repeated at two-yearly intervals. The questions on EXCALE are multiple choice items.[7] Students take the test in their last year of primary and secondary school. The Spanish test assesses three aspects: reading comprehension, reflection of language and writing. Based on their results for the reading section, students are classified into four categories:

1. *below basic* - students in this category show important deficiencies in knowledge and the use of strategies and skills. This may limit their ability to keep up with progress
2. *basic* - students possess the minimum level of knowledge, strategy use and skills
3. *medium* - students have an adequate level of knowledge of the curricula. They show a good performance in the use of strategies and skills
4. *advanced* - students have a superior level of knowledge and utilise strategies and skills effectively. They are expected to profit fully from the curriculum

Note that this classification specifically defines the use of readings strategies and skills as an important criterion for performance assessment.

Figure 2.1 shows the percentage of students in each of these categories for the assessments conducted between 2005 and 2007. Two results stand out: first, the vast majority of students are classified as having basic or below basic reading skills. This is a critical problem because reading is a basic skill for further studies and a predictor of later academic achievement (Lehto et al., 2001; Hakkarainen, Holopainen and Savolainen, 2013). If the students are not fully prepared, this might cause high dropout rates or repetition of the school year. Second, it appears that, as students progress through the educational system, the proportion of learners showing a basic

[7] The scores range from 200 to 800 points, with a mean of 500 and a standard deviation of 110.

or below basic reading competency increases. Whereas 54% of students show basic or below basic reading competency while at kindergarten, this increases to 71% for those at secondary school.

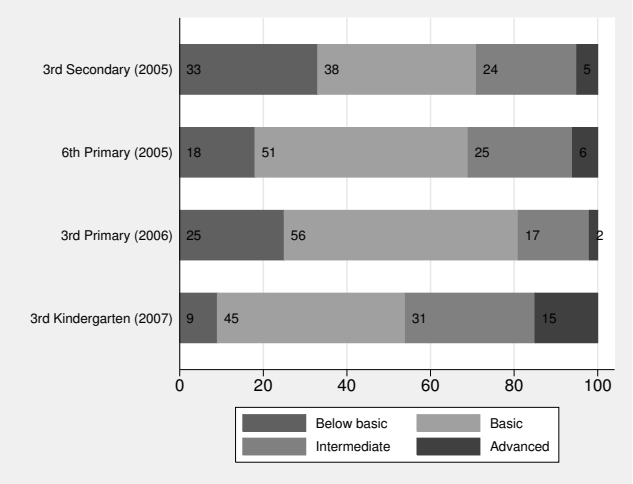


Figure 2.1.: Percentage of students in the EXCALE reading test categories (2005 - 2007).
(Source: Martinez Rizo (2008))

These results might indicate problems with the design of the syllabi or the quality of the education (e.g. inefficient teaching methods, large class sizes, etc.). In order to address the problem, the SEP has implemented several new programmes, such as the National Reading Programme (PNL - Programa National de Lectura).[8] However, the results of assessments in 2006 to 2009 showed a similar pattern as before, indicating that little progress had been made (Figure 2.2).

Two further assessments were implemented by the Mexican government. The National Standards Test (Pruebas de Estandares Nacionales) was administered first in 2000 and, in an attempt to analyse educational trends, the INEE repeated it in 2005. A comparison of the results revealed that primary school students had improved from 2000 to 2005, although not greatly. In contrast, the improvement in secondary school students, which is the focus of the present study, was small and statistically insignificant (Martinez Rizo, 2008). Although the report acknowledges improvement in primary school students, it deems this insufficient when compared with countries with similar or lower resources (e.g. Cuba).

[8]The National Reading Programme intends to improve and stimulate students reading competencies in basic education. The PNL has taken several measures, such as updating the curriculum, delivering specialised training for teachers and providing schools with books.

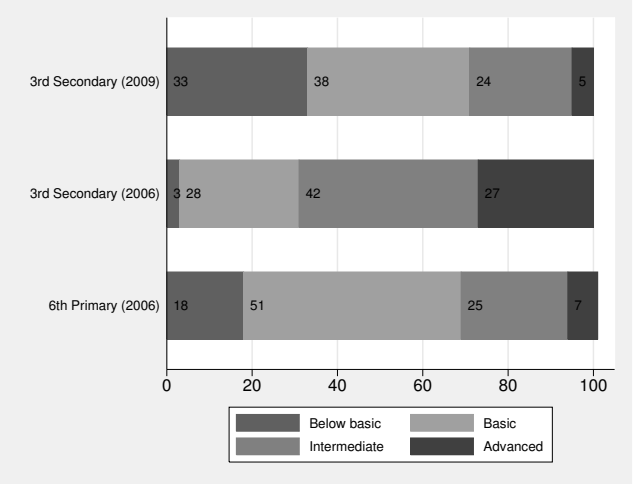


Figure 2.2.: Percentage of students in the EXCALE reading test categories (2006 - 2009).
(Source: Martinez Rizo (2008))

2.2.2. ENLACE

Mexican primary and secondary school students have been taking the National Evaluation of Academic Achievement in Academic Centres (ENLACE) test since 2006. The results have been used by Mexican schools to market themselves to current and future parents. ENLACE classifies students reading performance into four achievement groups: insufficient, elemental, good and excellent. Table 2.2 and Figure 2.3 show the percentage of students performing at each level. These results show a relatively stable pattern: first, over the years, approximately 80% of the students have performed at insufficient or elementary levels. Second, although there have been some changes in the percentage of students performing at each level, there has been little change in the number of students performing at good or excellent levels.

	Levels of achievement			
Year	Insufficient	Elemental	Good	Excellent
2006	40.7%	44.6%	14.0%	0.7%
2007	36.3%	44.8%	17.9%	1.0%
2008	32.9%	49.2%	17.1%	0.8%
2009	31.7%	49.5%	18.0%	0.8%
2010	39.7%	42.7%	16.6%	1.0%
2011	40.1%	42.8%	16.1%	1.0%
2012	37.9%	41.3%	19.0%	1.7%
2013	37.4%	42.9%	18.0%	1.7%

Table 2.2.: Percentage of students performing at each level of achievement in ENLACE

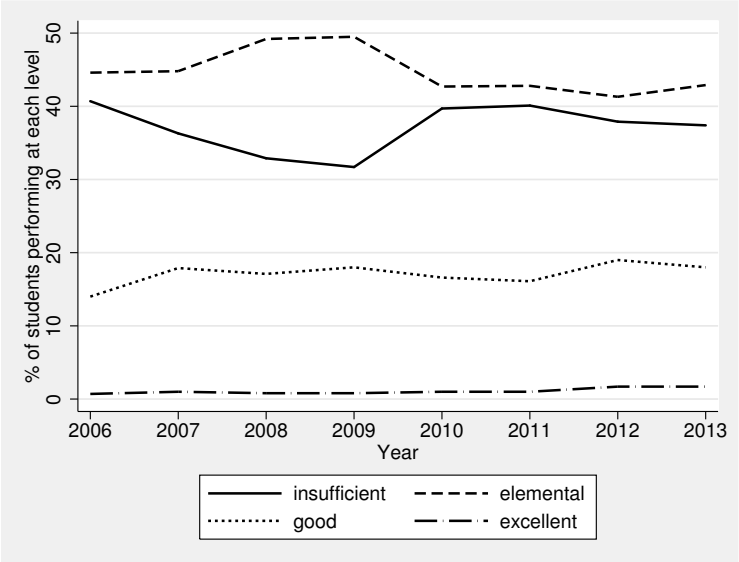


Figure 2.3.: Trend of students' performance in reading comprehension in ENLACE

2.3. International Evaluations: PISA

2.3.1. What does PISA evaluate?

The OECD plays an important role in the assessment of the performance of educational systems. PISA assessments cover the areas of science, mathematics and reading, and are carried out every three years. Every cycle emphasises a different area and reading has been the main focus of PISA in the years 2000 and 2009 (Table 2.3). However, all of the assessments contain tests designed to assess reading performance, so that country profiles can be compared longitudinally.

In PISA, reading is defined as *"the ability to understand, use and reflect on written texts in order to achieve one's goals"* (OECD, 2001, p. 21). To measure this understanding, PISA evaluates reading, taking into consideration three broad 'aspects' (or reading skills): a) access to and retrieval of information, b) integration and interpretation of information, and c) reflection on and evaluation of the content and form of texts. The abilities necessary to achieve each aspect are explained in detail in Section 2.3.3. However, before turning to this, it is important to explain the range of texts used in PISA, as different kinds of texts require different reading skills and the PISA texts form the basis of the reading comprehension test used in this study.

2.3.2. Texts in PISA

Texts in PISA are categorised according to their format, type, and the aspects required in order to comprehend them. Based on their format, PISA distinguishes between continuous and non-continuous texts. Continuous texts are defined as those composed of sentences organised into paragraphs. Examples of these are books,

Year	Emphasised subject
2000	Reading
2003	Mathematics
2006	Science
2009	Reading
2012	Mathematics
2015	Science

Table 2.3.: Schedule for PISA assessments and emphasised areas

newspapers, essays, and short stories. Non-continuous texts are frequently organised in matrix format or as a combination of objects, such as lists, tables, graphs, or diagrams. PISA assessments normally include about two-thirds of tasks addressing continuous texts and one-third of tasks addressing non-continuous texts (OECD, 2001). The assessment in 2009 included an additional classification: mixed and multiple texts. Mixed texts, found in magazines, reports, and electronic media, assess the skill to integrate information from different formats and texts; they contain tables or graphs to provide a better explanation or clearer example of what has been said in prose. Multiple texts are individually-written texts that, together, lead to a different conclusion than would be the case if they were read in isolation. Tests based on these texts make it possible to identify the readers' ability to form a relationship or contraposition. PISA 2009 also introduced the use of electronic texts; however, this area and the processes involved in the comprehension of such formats are beyond the scope of this study. The classification of text type is based on an adaptation of Werlich's classification (Werlich, 1976) and texts are defined as a description, narration, exposition, argumentation, instruction, or transaction. The aspects evaluated by the PISA test will now be explained in more detail.

2.3.3. Aspects of reading evaluated in PISA

Aspects are defined as *"mental strategies, approaches or purposes that readers use to negotiate their way into, around and between texts"* (OECD, 2009, p. 34). The five basic aspects to achieve comprehension are deemed to be:

1. retrieving information
2. forming a broad understanding
3. developing an interpretation
4. reflecting on and evaluating the content of a text
5. reflecting on and evaluating the form of a text

In order to operationalise these aspects and design appropriate test questions, PISA organises these aspects into three broad categories:

a) *Access and retrieve* - tasks in this category require the identification of separate pieces of information within a text (i.e. locating details, facts, etc.).

b) *Integrate and interpret* -tasks in this category assess the reader's ability to form relationships within a text and sometimes between texts. The reader creates a broad understanding of the text and makes inferences when the relationships are not clearly stated. Readers *"make internal sense of the text"* (OECD, 2009, p. 36). Summarisation and comparison are good examples of this category.

c) *Reflect and evaluate* - tasks in this category require readers to use their external knowledge in order to create an opinion of the text based on arguments drawn from their previous knowledge. They require them to be objective in evaluating the quality of texts. Examples of these tasks include evaluating the usefulness of information as well as the author's purposes.

Table 2.4 shows the distribution of the aspects in the 2009 assessment. The composition of earlier and later assessments is comparable.

Aspect	% of tasks
Access and retrieve	25
Integrate and interpret	50
Reflect and evaluate	25

Table 2.4.: Percentage of tasks by aspect in PISA 2009

2.3.4. Comparisons with the OECD average

2.3.4.1. Overall performance

Mexico has participated in PISA since 2000. The results show that, between 2000 and 2003, there was a significant decrease in the students' performance (Table 2.5). Subsequent years saw a steady but relatively small improvement and the difference between the 2012 and 2000 scores are statistically significant. The observed improvement has been explained to be a result of the higher socioeconomic status among the 2006 cohort (Martinez Rizo, 2008).

Due to the low ranking of Mexican students in 2000, the SEP launched programmes to foster reading comprehension (see Section 2.5). Although there have been some improvements, Mexico has always performed below the OECD average and has improved only slightly. This suggests that these programmes have not made any noticeable difference.

	2000	2003	2006	2009	2012
Mexico	422	400	410	425	424
OECD average	501	497	495	499	500

Table 2.5.: Results of Mexican students in reading in PISA (2000 - 2012). (Source: OECD PISA database (OECD, 2013))

2.3.4.2. Reading proficiency levels

PISA classifies students' performance into five[9] levels of proficiency (see Table 2.6). Students are classified into each level based on the score for the reading tests.

Proficiency level	Score points
5	$\geq$626
4	553 to 625
3	481 to 552
2	408 to 480
1	$\leq$407

Table 2.6.: Definition of PISA reading proficiency levels

Students in level five are considered 'top performers', while students' performance classified as level one is indicative of deficiencies in the use of reading literacy skills. Test items for each of the three aspects evaluated in PISA are classified according to these proficiency levels. For example, accessing and retrieving information can range from finding a telephone number (proficiency level 1) to finding synonymous information or discriminating between two similar pieces of information (proficiency level five). Over 40% of Mexican students fail to reach level two, which is defined by PISA as the minimum level required to perform adequately in developed economies (Table 2.7).

2.3.5. Comparison with other Latin American countries

One of the main aims of regular PISA assessments is to allow comparisons across and within countries in terms of levels of reading ability. Comparisons across developed

[9]This has subsequently been extended to eight levels by splitting level five into two levels, and level one into three levels. However, to ensure comparability over time, the initial classification is used here.

Level	2000	2003	2006	2009	2012
Mexico					
5	0.9%	0.5%	0.6%	0.4%	0.4%
2 to 4	55.1%	47.4%	52.4%	59.5%	58.6%
1	44.2%	52.0%	47.0%	40.1%	41.1%
OECD average					
5	9.5%	8.3%	5.8%	12.8%	8.4%
2 to 4	72.7%	72.8%	63.3%	73.0%	73.6%
1	17.9%	19.1%	31.0%	14.3%	18.0%

Table 2.7.: Performance of Mexican students grouped by PISA reading proficiency levels (2000 - 2012)
(Source: OECD PISA database (OECD, 2013))

and developing countries might seem unfair due to the differences in resources invested in education. Comparisons across a set of more homogeneous countries (e.g. all developing countries), however, are a more useful way of measuring achievement in educational policy and sharing experiences between countries with similar characteristics, problems and interests. In what follows, Mexican students' performance is compared with that of students from other Spanish speaking Latin-American countries who participate in the PISA assessments. It is worth noting that two factors limit the usefulness of this comparison. First, only six Spanish speaking countries in Latin America have participated in the PISA evaluations so far: Argentina, Chile, Colombia, Mexico, Peru, and Uruguay.[10] Second, only Mexico has participated in all five evaluations.

As stated above (see Table 2.5), Mexico declined in terms of reading performance between 2000 and 2003. The results for 2006 show a better performance than those for 2003, but are still lower than those for 2000. A comparison between the results for 2000 and 2009, when the tests focussed on reading, show that Mexico improved by 3 overall points. Comparisons with other countries show that Argentina also declined between 2000 and 2006 whereas Chile improved by 32 points (see Table 2.8). Comparing 2000 and 2009, Chile and Peru improved their reading performance by more than 30 score points. In the light of these changes in other

[10]Brazil, which has also been assessed by PISA, is not included because it is not a Spanish-speaking country.

countries, the 3 point improvement in Mexico must be judged as small.

Country	2000	2003	2006	2009	2012
Argentina	418		374	398	396
Chile	410		442	449	441
Colombia			385	413	403
Peru	327			370	384
Uruguay		434	413	426	411
Mexico	422	400	410	425	424

Table 2.8.: Trend in reading performance across Latin American countries (Source: OECD PISA database (OECD, 2013))

Comparisons of performance levels over time make it possible to identify which groups of students showed improvements. Table 2.9 shows the percentage of students performing below level 2 (i.e. the basic proficiency level) and those performing at level 5 or above in the years 2000 and 2012. Colombia and Uruguay are not included in this table as they did not participate in the first evaluation in 2000.

In 2000, around 44% of Mexican students performed at level one ('inadequate') and less than 1% performed at level five ('top performers'). In 2012, there was a reduction in the percentage of students performing at level one (-3.1%). Unfortunately, there was also a reduction in the percentage of students performing at level five (-0.5%). This suggests that the distribution of students' performance became more concentrated in the middle levels, i.e. levels 2-4. Argentina, with a similar percentage of students classified as performing inadequately in 2000, saw an increase in this proportion of just short of 10 percentage points over the 12-year period. In contrast, Chile, also similar at baseline, saw a decrease of just above 15 percentage points, which is substantial compared to the reduction of 3.1 percentage points in Mexico. Peru also saw a reduction of approximately 20 percentage points in the proportion of students reading inadequately but this is not directly comparable to the Mexican situation as Peru started from a substantially lower baseline level. Regarding the proportions of the top performers (level five), only Chile and Peru experienced a slight increase. Based on these comparisons, and assuming that the observed changes are statistically significant, one might ask which actions these countries have taken to improve their performance, and how their policies and programmes have been designed to induce these changes. However, this is beyond the scope of the current

	2000		2012		Difference (2012 - 2000)	
Country	Level 1	Level 5	Level 1	Level 5	Level 1	Level 5
Argentina	43.9 %	1.7%	53.5%	0.6%	9.6%	-1.1%
Chile	48.2%	0.5%	33.0%	0.6%	-15.2%	0.1%
Peru	79.6%	0.1%	59.9%	0.5%	-19.7%	0.4%
Mexico	44.2%	0.9%	41.1%	0.4%	-3.1%	-0.5%
OECD average	17.9%	9.5%	18.0%	8.4%	0.1%	-1.1%

Table 2.9.: Percentage of students in Latin American countries performing at levels 1 or 5
(Source: OECD PISA database (OECD, 2013))

study.

2.4. International evaluations: Others

Several other international evaluations have been carried out. The Latin American and Caribbean Regional Office of UNESCO promoted a programme to compare the quality of education among the countries in this region. The Latin American Laboratory for the Assessment of Quality in Education (LLECE in Spanish) allowed a comparison of primary education amongst 13 countries in the area. The first study (Primer Estudio Regional Comparativo y Explicativo, PERCE) evaluated 3rd and 4th grade primary school students. It focussed on two areas: the national language (Spanish or Portuguese) and Mathematics.[11] The results placed Mexico above many countries with similar or worse socio-economic characteristics. The second study (Segundo Estudio Regional Comparativo y Explicativo, SERCE), carried out in 2006, included Science. SERCE was taken by students in the 3rd and 6th grade of primary school in 16 countries.[12] The results for the reading part of both tests are shown in Appendix A.2. Although Mexican students' performance in these tests is better than that of students in other developing countries, there is still concern about why Mexico performs below countries with serious economic problems (i.e. Cuba).

[11] The mean achievement score of PERCE is 250 with a standard deviation of 50.
[12] The mean achievement score of SERCE is 500 with a standard deviation of 100.

2.5. National reading policies

As a result of Mexican students' low performance, the Mexican government launched the National Reading Programme (PNL) in 2001. Its main objective is to improve reading skills by giving support to schools and provide necessary resources (SEP, 2012). The programme is implemented in schools of basic education (i.e. kindergarten, primary and secondary stages). It involves the continuous updating of the curricula and provision of resources in school libraries as well as a schedule with detailed activities for stakeholders in the educational system (students, teachers, head teachers, parents, etc.).

The schedule provides guidelines about the use of specific activities to develop a plan of action/work. It also helps the institutions to fulfill the objectives of the national strategy called '11+5: Actions to integrate a community of readers and writers'. This strategy involves three main areas of action, which aim to: 1) define the roles for each member of the educational system, 2) ensure the correct implementation and use of the school library, including the election of the reading and library committees, and 3) promote five permanent reading activities. The first two areas of action provide the logistics and processes to follow in order to provide schools with resources appropriate for developing reading activities (i.e. a school library and classroom library). The third area of action, managed by the teachers, might have a direct impact on the students as it is designed to improve their reading skills. In this area, the teachers are responsible for the implementation of the following activities during the academic year: 1) reading aloud, for 15 minutes on a daily basis; 2) reading one book per month, including a session at the end of the month to share ideas about the book; 3) reading at home, 20 minutes of reading with parents who complete a report; 4) inviting reading guests to share stories with students; and 5) keeping a record of the number of books read by the students. In order to keep track of the students' improvement, the PNL included an extra activity in academic year 2012-2013. The teachers were asked to submit a monthly report of the number of words per minute read by each student during the reading aloud sessions.

Although the programme places a strong emphasis on resources and practice, it does not include activities intended to provide students with better guidance about comprehension. Its main focus is the availability of resources and facilities as well as the administration and logistics of the libraries. It fails to provide details about how the learners themselves might achieve better text comprehension and how the teachers might help them to do this. The current study aimed partially to address

that need, for one age group of pupils.

3. Literature Review

Reading allows individuals to acquire knowledge and obtain information; a crucial skill which facilitates the integration of individuals into society. Reading involves the interaction of several independent processes at different levels. The degree of involvement and interaction of these processes varies and depends on the reader's specific goals. The complexity of reading has led to the development of multiple research areas such as research into i) the cognitive processes underlying reading and the way they operate simultaneously in a harmonious way (Rumelhart, 1994; Stanovich, 1980), ii) comprehension development (Paris and Hamilton, 2008; Pressley, 2000), iii) reading strategies used by skilled and less skilled readers (Garner and Kraus, 1981; Pressley and Afflerbach, 1995), iv) methods to enhance comprehension, specifically comprehension instruction (Block, 2004; Duffy and Roehler, 1989; Paris, Cross and Lipson, 1984), and v) intrinsic and extrinsic factors affecting reading performance (Samuels and Kamil, 1984).

In the last decades reading research has moved out of the laboratory and into the classroom. Evidence from observational studies and field experiments has led to the integration of new reading approaches into the curricula, particularly in developed, English speaking countries such as the United States and the United Kingdom. Also, as a result of globalisation and advances in technology, interest in the processes involved while reading texts embedded in virtual environments (e.g. internet webpages) and comprehension instruction in multicultural communities has increased (Raphael et al., 2008) [1].

The aim of this chapter is to provide the methodological foundation for the applied phase of work and place this thesis into the context of existing studies on direct teaching of reading comprehension strategies. This chapter is therefore not an exhaustive review of previous research in the field of reading. Instead, it offers a general overview of the processes involved in reading and the different researched areas relevant to this study. The main focus of this chapter is on reading comprehension strategies

[1] These issues were deemed outside the scope of this literature review and not important for the quasi-experimental study. They are therefore not further considered

and the way they are taught. Emphasis is given to strands of literature that informed the aims and methodology employed in the current study of a taught intervention of reading comprehension strategies. The chapter is divided into three parts: The first part addresses the reading processes and comprehension development. The second part reviews learning theories and shows how learning strategies fit within the theory of cognitivism. Finally, teaching approaches to reading comprehension and applied studies of direct teaching of reading comprehension strategies are reviewed.

3.1. Reading and comprehension

3.1.1. Reading processes

Reading involves simultaneous mental operations (i.e. cognitive processes) that allow readers to decode printed symbols in order to comprehend the meaning of a text. Multiple definitions have been developed in the literature, yet no single and universally accepted definition of reading exists. Some authors describe reading as the cognitive processes carried out to access author's ideas (Rayner and Pollatsek, 1989). Others, such as Grabe (2008), develop the concept of 'kinds of readings' based on readers' levels and abilities as well as their purpose of reading.

The multiple cognitive processes involved in reading are not activate at all times. The activation of specific processes depends primarily on the purpose of reading. Reading to search for information and reading for quick understanding, for instance, do not require high levels of cognitive processes; reading to learn, to integrate information, and to evaluate or critique, however, require high levels of specificity and detail. Reading is regarded as *"a unified process that is adjusted flexibly in response to reader purpose, reader proficiency level, and possible contextual constraints"* (Grabe, 2008, p13). Therefore, different cognitive processes are triggered depending on the type of reading activity. Overall, reading depends on rapid, efficient, comprehensive, interactive, strategic, flexible, purposeful, discerning, linguistic and information-gathering processes (Grabe, 2008). The main purpose of reading, regardless of the 'kinds of reading', is reading to achieve comprehension of a text.

The different cognitive processes that are involved in reading have attracted considerable interest of researchers in both education and psychology alike. Psychologists have focused on the mental processes involved in reading and on the development of comprehension. These mental processes are typically categorised into two types: lower-level and higher-level processes (Cain, Oakhill and Bryant, 2004; J. Field,

2003). Lower-level processes are skills that are expected to become automatized during early education and are henceforward carried out unconsciously. In contrast, higher-level processes are in constant development over the course the reader's life. Although theoretical models describe these types of reading processes to occur simultaneously, this classification allows a better understanding of how inefficiencies in lower-level processes might affect the efficiency and development of higher-level processes. Following this classification, the next section explains the cognitive processes working under 'normal' reading conditions, i.e. processes used by skilled readers of orthographic alphabetic systems.

3.1.1.1. Lower-level processes

Lower-level processes constitute the decoding of phonological codes (i.e. word recognition), accessing of meaning, syntactic parsing and semantic proposition formation. The reading process of skilled readers begins with visual recognition and deciphering of printed symbols (i.e. words). Letters in isolation are recognised and identified as patterns that possess unique characteristics such as spelling, pronunciation, and meaning (Carroll and Snowling, 2004). Thus, letters represent not only symbols but also sounds that are translated into meaning. Some sequences are more likely to appear than others (i.e. familiar words). The orthography of the language is learned by continuous exposure to these patterns resulting in the automatic identification of familiar words or words following regular spelling. Word recognition normally occurs rapidly and accurately. Readers that struggle to recognise words find it difficult to gather the meaning of the text. Therefore, word recognition abilities at early age are generally regarded as one of the most important predictors of later reading ability (Perfetti, Landi and Oakhill, 2005; Perfetti, 2007). Inefficient word recognition presents a major obstacle for learners.

A critical factor for the efficiency of the lower-level processes is the working memory. Memory is divided into permanent records of our experiences (i.e. long-term memory)(LTM) and transient ones (i.e. working memory or short term memory (STM)). Working memory is a mental place where information is stored for brief periods of time (Baddeley, 1992). It supports phonological, orthographic, and morphological processes for word decoding by keeping the information active for a few seconds while processing operations take place. Its main drawback is its limited capacity. Theoretical models of reading suggest that word recognition problems might be cause by a limited capacity in the reader's working memory (Baddeley, 1979;

Daneman and Carpenter, 1980). Text comprehension and word recognition take up some capacity of the working memory (Kintsch and Van Dijk, 1978). Information is active in the working memory until it is transferred to and integrated into the long-term memory. Unskilled readers are thought to store large amounts of information in the working memory, thereby exhausting its capacity and leaving insufficient room for processing other information (Daneman and Carpenter, 1980). Automatisation of word recognition frees capacity in the working memory and allows the reader to carry out processes operating at higher levels. Extensive research in the area of word recognition has been carried out in order to develop and test approaches to improve students' word recognition (Abbott and Berninger, 1999; Samuels, 1988).

Word organisation is another important process that leads to comprehension. The parsing of syntactic information provides information about the role that a word plays in the context of other words (typically a sentence, i.e. noun, adjective, etc.) and how it co-indexes with other elements in the discourse (e.g. pronouns, connectives). After the words are recognised, the orthography (i.e. visual recognition of words) and lexical information (i.e. decoding of symbols) are combined with syntactic information (e.g. nouns, verbs) to understand the meaning of words. The meaning of a word depends on its role in the sentence and readers make use of context to choose between alternative interpretations when multiple meanings are possible (e.g. does the word 'chair' function as a noun or a verb in any particular context?). Meanings of individual words are collated into a collective and coherent interpretation of the sentence. Unskilled readers invest a great amount of time and effort to recognise words and to activate appropriate semantics. This extra work decreases the resources available to remember preceding words when combining different words into a sentence.

It has been argued that phonology serves as a backup system for visual recognition (Samuels and Kamil, 1984). The advantage of the phonological processor is the source of input, speech over visual, which provides a second source of information in instances where the symbols are poorly recognised or not at all familiar (Adams, 1990). Because a language is first learned orally during children's early life years, individuals learning in their first language already possess a substantial understanding of that language. They know its sound systems, distribution of elements, and sound to meaning connections before they first encounter printed symbols. The phonological processor is highly active when reading aloud and it contributes to accuracy and high-speed recognition as well as to the expansion of the on-line memory. This alternative source of inputs aids the reader in coping with unfamiliar words. Overall, the

coordinated work of the orthography and phonology helps to overcome confusions and to compensate for difficulties.

The formation of a mental model is the last lower-level process. Once the meaning of individual words is known, readers must contextualize the meaning of the sentence to understand the role of the sentence within the text. This is known as the semantic information (Rayner and Pollatsek, 1989). Semantic information helps to access the meaning of a sentence through the use of propositions and the subsequent creation of a mental model (Perfetti, Landi and Oakhill, 2005). The propositions help the reader to identify and subsequently exclude less important words from a sentence. As a result the reader obtains a representation of the meaning of the sentence. Propositions of different sentences are combined to form a mental model of the text (Grabe, 2008), or, what Kintsch and Van Dijk (1978) called, the semantic structure.

Different models of the reading process have been proposed. Some of these models assume that the various lower processes are carried out simultaneously (e.g. Samuels and Kamil, 1984). The interaction of the semantic and syntactic information, for instance, might facilitate the spread of word activation mechanisms. Although semantic and syntactic information become available after word recognition, it has been assumed that words that are recognised (i.e. lexical access) activate the semantic of the neighbouring words in the sentence as readers are sensitive to what is likely to follow using distribution probabilities (Grabe, 2008).

3.1.1.2. Higher-level processes

Higher-level processes help readers to integrate the understood sentences into the overall interpretation of the text. Several reading models have been proposed to explain the operational processes involved in the higher-level reading processes (Samuels and Kamil, 1984). The models of Just and Carpenter (1980) and Kintsch and Van Dijk (1978), for instance, focussed mainly on comprehension. Following experimental research, these models have been modified and adapted but their core assumptions remain unchanged. They assume that interpretation of texts includes: a) understanding of a sentence within a text, b) retrieval and integration of information presented in different parts of the text, c) integration of knowledge extraneous to the text (e.g. previous knowledge), and d) inferential activities.

One of these models suggests that readers create a model in their minds formed by the association of *anaphors, propositions and inferences* (J. Field, 2003). *Anaphors* are words or phrases that refer back to earlier parts of the text and help readers to

differentiate between 'new' and 'given' information within a sentence or group of sentences (Clark and Haviland, 1977; Harley, 2001). *Propositions* are retained in the STM while connections are formed. Successful comprehension of the information can be achieved through two different mechanisms. First, the reader tries to match the new information represented by propositions with those propositions previously stored in the STM. If the information matches, the text is considered coherent. The second mechanism is activated when the information does not match and the reader searches for a match in his LTM (i.e. prior knowledge). This highlights the importance of prior knowledge for comprehension (Kintsch and Van Dijk, 1978). If none of the routes provide a coherent interpretation of the text, *inferences* are formed. Inferences are new propositions created by means of information stored in the readers' schemata. In short, appropriate identification of new information (propositions), its integration with previous knowledge, as well as the creation of inferences are vital steps to giving coherence to a text. Readers' ability to make inferences helps them to integrate different ideas from the text. Conversely, failure to construct inferences or a poor knowledge of the topic might result in poor comprehension.

Interlinked propositions form the semantic structure of the text which consist of the microstructure and the macrostructure. The microstructure represents local events, i.e. the individual propositions as well as their relationships. The macrostructure is formed by hierarchically arranged units of the microstructure linked to a pre-existing 'schema' to create a global representation of the text (Harley, 2001). The omission of unimportant propositions and the addition of inferences allow the transition from microstructure to macrostructure.

An alternative model states that readers possess 'schemas' or ideas, which are representations of situations or events in the real world. Schemata (group of schemas) facilitate comprehension and allow accurate recollection of the essence of a text, which, in turn, gives meaning to the text (J. Field, 2003). The association of micro- and macrostructure provides the reader with only a basic understanding of the text (Harley, 2001). In order to gain a deeper understanding, it is necessary to combine the text-base formation (micro- and macrostructure) with readers' prior knowledge and their goals to form the situational model (J. Field, 2003; Kintsch and Rawson, 2005). The situational (or mental) model is a propositional representation of the information that allows readers to make representations of the different relationships between sentences available within the text (Van Dijk and Kintsch, 1983).

The last component commonly mentioned as part of the higher-level processes is the use of strategies, monitoring, inferences and goals. These elements are the

primary focus of this thesis and are discussed separately in this chapter.

3.1.2. Comprehension

"Comprehension is the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (RAND Reading Study Group, 2001, p. 11). It is often treated as synonymous of reading. In the previous sections it was described as the final product of the reading process as well as the main purpose of reading. However, it is best explained as the 'product of reading' because comprehension develops through practice and time (Alderson, 2000). Scholars agree that comprehension of a text can be achieved through the interaction of the reader with the text at different levels. Gray, for example, proposed three levels of difficulty in reading: *"Reading can be performed by reading 'the lines', 'between the lines', or 'beyond the lines'"* (Gray, 1960 in Alderson, 2000, p. 7). That is, readers are able to achieve a literal understanding of the text either through decoding, by making inferences, i.e. gaining an understanding of meaning not directly stated in text by employing prior knowledge, or through the critical analysis of the text, i.e. inferring the main implication of the text; again using prior knowledge.

Success at the first level ('the lines') is achieved by mastering word decoding and lower-level reading processes. An appropriate word decoding (i.e. grapheme phoneme processing) allows automatic understanding of word meaning, which frees space in readers' short-term memory and facilitates the operation of higher-level processes (see above). Some researchers have argued that training in decoding and word comprehension can improve general comprehension (Juel, Griffith and Gough, 1986; Yuill and Oakhill, 1991). Therefore, it could be advisable to train students in decoding words until, after sufficient practice, this process becomes a skill, i.e. an automated process. However, whether word decoding improves general comprehension is still being debated: while some studies found that training in decoding and word comprehension improves comprehension (A. Cunningham, 1990; Lundberg, Frost and Petersen, 1988; Tan and Nicholson, 1997) others do not consider it as an effective approach (Yuill and Oakhill, 1988; Yuill and Oakhill, 1991). Yet, despite controversy, grapheme-phoneme theory has proved greatly influential and is now being followed in many education system (Carroll, Bowyer-Crane et al., 2011; Hulme and Snowling, 2005).

Further to word decoding, research suggests that an extensive vocabulary benefits comprehension (Beck, Perfetti and McKeown, 1982). Hence, experiments have been

designed to determine the best approach to teach vocabulary (Blachowicz and Fisher, 2009; Nagy, 1988; Nation and Newton, 1997). The results of these studies generally suggest that most of the vocabulary is learnt while reading it in a text, i.e. vocabulary is best learned in the context in which it was first encountered and remembered. Because readers will always find new words within texts, some vocabulary research focuses on finding ways to help readers to understand unknown words by making use of context.

Success at the second ('between the lines') and third levels ('beyond the lines') can be achieved by adequate development of comprehension at word-level and above word-level (Pressley, 2000; J. Field, 2003). Above word-level, squemata and the macrostructure theory suggest that background knowledge affects comprehension because good comprehension derives from an appropriate interaction of reader's previous knowledge with the ideas that are represented in the text. Information of previous knowledge is stored in related structures, which are easily accessed by readers while reading. Readers retrieve information (i.e. schemata) and make use of it to fill in missing details. Schemata is a preconception of the situation in the world that allow the reader to make inferences by following top-down processes; that is, going from the general concept to its particulars. Hence, poor prior knowledge might indirectly contribute to poor comprehension through the lack of schemata.

Informed by previous research and an understanding of how individual processes work, attempts have been made to explain the reading process by developing reading theories. The simple view of reading, for example, is thought as the combination of two components: a) decoding and b) linguistic comprehension (Hoover and Gough, 1990). While the two components are necessary to achieve comprehension, the relationship of these components changes along the cognitive stages of the reader. Decoding correlates strongly with students' comprehension in their early years whereas linguistic comprehension is highly correlated with comprehension in later years. Decoding is regarded as a difficult skill. However, languages with similar or transparent grapheme-phoneme correspondence (e.g. Spanish) makes decoding easier than in less transparent languages such as English. This would imply that, at least in Spanish speakers, enhancing linguistic comprehension is of higher importance. That is, parsing, bridging, and discourse building should be emphasised when designing reading instruction to improve students' understanding of texts.

The construction-integration model (Kintsch, 1988) involves the combination of a bottom-up process where new concepts are combined to similar concepts previously stored in the long-term memory to create a representation of the world. When these

new concepts are not in line with previous information, the reader infers possible alternatives based on their prior knowledge to integrate these concepts.

A range of different factors may affect comprehension development. According to Pressley (2000) the three most important factors that are primarily responsible for comprehension development are word-level skills, background knowledge, and comprehension strategies. Word-level skills and use of prior knowledge are regarded as automatic and unconscious lower-level processes. Comprehension strategies, however, are conscious processes that can be controlled (and monitored; see before) to affect comprehension development. Therefore, strategy instruction might serve as a tool for improving comprehension.

Strategies for comprehension development are thought to play such a crucial role in the development of reading comprehension that a large amount of research activity has been dedicated to it. Studies have focussed on a wide range of aspects, such as definition and characteristics of reading comprehension strategies, description of their use, effectiveness, teaching techniques, etc. Given the importance that comprehension strategies play in the empirical work presented in this thesis, a detailed description of the literature on reading comprehension strategies is presented in Section 3.3.

A commonly overlooked factor affecting comprehension is the 'standards of coherence'. Comprehension involves mental representations of the text; however, the relationship between these images needs to be coherent for the reader to make sense of them. Text might include referential, causal, temporal, or logical coherence to mention some. When assessing the information in a given text, the reader assesses if the coherence in the text is adequate based on his own standards of coherence (Van den Broek et al., 2011). These standards are his/her own criteria against which the new information is compared. The reader will apply different criteria for different texts and reading goals. Readers' criteria will determine their desired level of understanding. Therefore, from the available reading strategies readers will choose the one that is adequate for their reading needs as informed by their standards of coherence.

3.2. Learning theories and learning strategies

3.2.1. A brief overview of learning theories

Learners learn in different ways. Mayer (1988) suggests three learning theories to explain how learners acquire new knowledge: quantitative, qualitative and behavi-

oural [2]. The first two learning theories were pioneered primarily within the field of education. The quantitative theory focuses on the units (or quantity) of knowledge learned. Assessment of this aspect of learning is carried out in the form of recall tests which measure how much knowledge has been acquired (e.g. a vocabulary test). The qualitative theory of learning focuses on the type of information learnt. It recognises that attention is a selective process and that different information is needed for different learning situations. It can be measured through recall of information as well as by means of elaboration of inferences and transference of information.

Achievement of an effective learning process requires a balance between the qualitative and quantitative models of learning, which is the rationale behind training students in the use of different learning strategies. These theories could account for the role of the learning strategies in that the learning strategies help to make the previous processes explicit to students - i.e. change the nature of the learning (the qualitative model) and increase the amount of knowledge stored (the quantitative model). It is expected that with practice, the strategies may become automatic processes.

In addition to theories developed by educational researchers, psychologists also provided an explanation of how learners learn. 'Behaviourism' was one of the first theories about learning. For behaviourists, learning is the acquisition of a new behaviour (A. Pritchard, 2009; Schunk, 2000). However, the learning theory of behaviourism has been critizised because it does not describe the learning process itself and treats it as a black box instead. Students receive information but behaviourism does not attempt to describe how they process this information. Hence, the explanation of the learning process was deemed incomplete (Mayer, 1988).

On the contrary, cognitivism focuses on mental processes used while learning or solving problems (Schunk, 2000). Cognition covers the mental activities that allow the acquisition, storage, retrieval and use of knowledge. It describes the ability of the brain to think and solve problems. Constructivism influenced cognitivism because cognitivists perceive learning as the result of mental construction. In constructivism students are assumed to use their experience to make sense of the information they perceive in order to build their own understanding. Cognitivism strongly influenced the conceptualisation of learning strategies. Students are able to carry out cognitive processes when acquiring knowledge. For cognitivists, learning strategies were *"integrated sequences of procedures selected with a purpose in view"* (Nisbet

[2]Mayer (1988) does not specifically mention cognitivism, although it played an important role in learning theories developed in psychology (see below).

and Shucksmith, 1986, p. 6). Successful learners will have developed a range of strategies from which they are able to select one or many appropriate strategies for each situation (Nisbet and Shucksmith, 1986). In order to select the appropriate strategy, students need to be aware of what they are doing, as well as of their own learning style, i.e. strategy selection is not necessarily an automated process. In addition, students may need to monitor their learning to make appropriate decisions. *"Learning to manage the process of learning involves being aware of what one is doing, or being able to bring one's mental processes under conscious scrutiny and thus more effectively under control. The awareness of one's mental processes has been termed 'metacognition'"* (Nisbet and Shucksmith, 1986, p. 7).

Monitoring involves thinking about one's performance while reading. It helps to assess the progress made towards comprehension and to identify comprehension failures. Comprehension monitoring is an on-going activity that requires the reader to take an active role in verifying whether the information that is being processes in the text is consistent with his prior knowledge and his standards of coherence.

According to Paris and Myers (1981), comprehension monitoring involves evaluation, planning, and regulation. If a problem is identified during the evaluation stage, the reader is expected to make a plan in order to solve the problem - by using conscious processes or strategies. The plan is then performed, which regulates the behaviour. Activities such as periodic reviews of what has been read, self-interrogation for clarification, interpretation and prediction, and summarisation are activities that trigger comprehension monitoring.

More generally, cognitive psychologists place learning strategies within an information-processing theoretical model which includes an operative (or cognitive) function and an executive (or metacognitive) function (Mayer, 1988). The role of learning strategies is to make explicit what otherwise may occur without the learner being aware, or what may occur inefficiently during early stages of learning.

With these differences in mind, researchers differentiated learning strategies into three different kinds:

1. Cognitive strategies - related with individual learning tasks that deal with direct manipulation of the information.

2. Metacognitive strategies - dealing with thinking processes, the planning, monitoring and self-evaluation of the learning activity

3. Affective/social strategies - involving the influences of the environment in the

learning processes, that is, how peer interaction can influence a common goal in learning.

Theories of constructivism proposed an active construction of students' knowledge. Piaget and Vygotsky, both constructivists studying children development, supported learning through action. While Piaget focussed on inner motivation to connect new information to previous knowledge, Vygotsky stressed the importance of social interaction (Smith, 1997; Driscoll, 2000).

Piaget believed in active learning where children operate individually. That is, knowledge acquisition is a continuous process of self-construction. Knowledge is created as children develop and interact with the world surrounding them. Children's development will occur gradually according to their age and in line with, but never exceeding, their capabilities.

Piaget distinguished three types of knowledge: physical, logical-mathematical, and social knowledge. The first two types of knowledge are acquired by the individual alone, whereas the third one, the social knowledge, is influenced by the actions of or interactions with other people. Therefore the role of the teacher, in instruction based on Piaget's ideas, is that of a provider or organiser. The teacher should provide situations for the child to work and thus discover their own knowledge. Children's interactions with their peers are another important source of cognitive development.

Instruction following Piaget's ideas emphasises a child-centred educational philosophy. It also includes i) the learning environment, ii) the peer interactions and iii) the creation of conflicts. An appropriate learning environment should be created to support the activity of the children by encouraging them to initiate and complete their activities to create their own conceptions. Peer interactions help children to move from their self-centred thought. Instructional strategies are favoured that encourage peer teaching and social negotiation during problem solving.Providing children with conflicted situations helps them to become aware of inconsistencies in their thinking. When they solve this imbalance, children move to the next stage of development.

Vygotsky also believed in active learning but he emphasised the importance of performing learning activities within a social group. He believed in the presence of a 'more knowledgeable person' who will help to scaffold the knowledge through collaboration. He regarded development as an internalisation of a social experience: individual development could not be understood without reference to the social and cultural context. That is, the interaction of less able learners with more knowledgeable peers will result in development of the first ones. He focussed on

the mechanisms of development by emphasising the illuminating process, how the individual performs under certain conditions, rather than the product. This process provides a mediational view of development. That is, in higher forms of human behaviour, the individual actively modifies or mediates the stimulus situation as part of the responding process. Piaget also studied skill development and proposed the Zone of Proximal Development, which is the potential development of a child under adult guidance or in collaboration with more capable peers.

Finally, a third perspective was provided by Bruner (Bruner, 1990; Bruner, 2006; Smidt, 2011). For him, children should reach an endpoint by thinking and building based on their experience. He thought that humans respond to the environment. Unlike Piaget, who regarded development as a fixed sequence, Bruner believed in the invariant sequence of stages. He emphasised the interaction of the learner with other learners or the teacher in order to develop and move from one learning stage to the next one. That is, interpersonal interactions provide learning tools. In contrast to Vygotsky and Piaget, Bruner emphasised the role of pre-elaborated examples in order for children to reach pre-desired conceptions.

Reading approaches e.g. reciprocal teaching have benefited greatly from constructivists. One point has been the development of student-centred approaches. That is, either by requesting a desired result (i.e. summarisation) or by allowing readers to build their own knowledge (i.e. inferences) reading approaches seek to form independent learners capable of understanding a text and analyse its contents critically. In order to develop this critical thinking, readers compare new concepts with their previous experience to become aware of the inconsistencies of the text. Another point that reading approaches have adopted is the social interactions, e.g. through teacher modelling or peer reading. By modelling a specific situation, teachers provide the role of the more knowledgeable person and interact with the students to help them scaffold their knowledge. Also, peer interaction in discussions and exchange of information based on readings tasks allow exchanging experiences and foster collaboration in order to reach a consensus to form the new knowledge.

3.2.2. Learning strategies

The importance of reading comprehension strategies has been emphasized in the previous section. They have been described as problem-solving tools to overcome understanding problems. However, for the strategies to be acquired, it is not only necessary to teach them but for students to learn them and use them. Therefore, it is

important to review the different learning processes and their role in the acquisition of skills.

The main goal of education is to provide individuals with learning opportunities. In the past, educational activities primarily consisted of teacher-directed activities where students were told what to do and how to carry out different tasks. Emphasis was placed on the correctness of the outcome rather than on the means for achieving that outcome. However, it was observed that individuals who developed learning strategies benefited from their use. This led some educational researchers to develop a research agenda on learning strategies to improve the provision of education, and ultimately, students' learning from a more learner-centered perspective.

Several definitions of learning strategies have been developed depending on the area of study. A general definition that applies to several areas of study is that learning strategies are *"behaviours of a learner that are intended to influence how the learner processes information"* (Mayer, 1988, p. 11), i.e. learning strategies entail different means to achieve a learning goal (e.g. comprehension). Students' preferences and background might affect the choice and effectiveness of the different learning strategy and thus the learning process itself.

In order to understand how learners acquire learning strategies, it is useful to distinguish between two main kinds of knowledge: declarative knowledge and procedural knowledge (J. R. Anderson, 2009). Declarative knowledge is defined by J. R. Anderson (2009) as 'static' because it is merely the information stored in the LTM, or 'what we know about', such as facts and definitions. In contrast, procedural knowledge is 'the ability to apply the knowledge of rules to solve a problem', i.e. the ability to understand. It is 'what we know how to do' and requires continuous practice to acquire. Learning strategies fall naturally within the second type of knowledge.

J. R. Anderson (2009) explains the acquisition of learning strategies as a complex cognitive skill and proposes a learning system to explain how strategies are acquired. This system contains three stages that a learner needs to pass through in order to acquire a learning strategy: cognitive, associative, and autonomous. In the cognitive stage the learner receives instruction on how to perform the task (procedural knowledge). In the associative stage, the learner might detect errors in how he or she executes the task (i.e. procedure). With practice, these errors are eliminated and thus the connections between new and previous information are strengthened. The final stage, the autonomous, allows the learner to apply or execute the knowledge quickly and automatically, without making mistakes.

3.3. Reading comprehension research

Reading comprehension is a widely studied area. Several topics have derived from its study e.g. comprehension development, reading instruction, etc. However, the review of studies in the following section follows a chronological structure rather than a thematic one. Although the reading strategies have been studied over time, the review of these studies shows a common pattern of methodological characteristics that have evolved over time. Therefore, based on these methodological characteristics, the studies were grouped following a chronological order.

This is not an exhaustive review but rather an inclusive one, which incorporates studies that have proven to be influential for this area of research. The parameters of the review were set based on the characteristics of the intervention. It was decided to review interventional studies teaching at least one reading strategy in order to enhance comprehension. Because this was a long-term project, the literature was searched up to 2011, which was the year when the intervention took place. The literature was reviewed as follows: first, a key word search including the words 'reading comprehension', 'direct teaching', and 'reading strategies' was run using Linguistics and Language Behaviour Abstract (LLBA), PsycINFO, and Dissertations and Theses databases. Relevant studies were retrieved and the references of these studies were scanned for further studies. In addition, previous reviews of studies were tracked, retrieved and reviewed. Studies reporting improvement of comprehension and those that had been replicated were retained. These papers allowed the creation of a list of strategies and methodological characteristics.

3.3.1. Reading comprehension instruction

3.3.1.1. Development of reading comprehension instruction over time

Comprehension instruction has been influenced by innovative approaches to learning and research on human behaviour. The development of cognitive psychology in particular has contributed greatly to the study of reading comprehension. Also, classroom instruction has been widely studied in an attempt to identify the most effective approaches to enhance students' reading performance (Pearson, 2008). Up until the 1960s, repeated practice of decoding and fluency were regarded as best practice (Mathews, 1966). Oral reading, focused on accuracy and expressive fluency, was at the core of classroom instruction as a means for assessment.

Between the 1930s and 1975, basal programs and students' materials (e.g. work-

books) were used as tools to facilitate the transformation of reading into a skill but without a detailed outline of specific skills (Pearson, 2008). Workbooks were considered especially useful and they were frequently employed to practise and enhance reading skills. Students' comprehension was practised in class through discussions led by teachers. It was thought that the practice of answering questions would lead to understanding. Also, the adoption of multiple-choice tests and standardised texts allowed rapid assessment of students' levels of literacy. Therefore, posing questions was particularly popular among teachers and formed the main tool used to guide students in conversations and post-reading discussions.

The development of new cognitive theories and the work carried out by Durkin (1978) in the late 1970s modified the perception of comprehension and reading, and marked the beginning of intense reading comprehension research. Durkin analysed comprehension instruction delivered in middle-school classrooms. Based on class observations and syllabi analysis, Durkin concluded that students' comprehension instruction mainly consisted of learning words and phrases, filling in work pages, answering questions, and taking tests. Little time was dedicated to offering advice about how to understand and comprehend texts (Pearson and Gallagher, 1983). As a result of Durkin's research, researchers began developing new approaches to provide students with adequate tools to enhance their understanding.

First, researchers focused their efforts on the study of basic reading processes. Structural schema theory, students' prior knowledge to reading, and the processes activated by characteristics of text structure were of special interest. Schema theory perceives the reader as a builder who processes the language and the information (Pearson, 2008). Instruction following this theory included elements from students' prior knowledge and cultural backgrounds into lessons, so that the combination of the students' own knowledge with the text produced a new schema. Motivated by these ideas, several studies were designed to test the hypothesis that prior knowledge was a good predictor of comprehension (Baldwin, Peleg-Bruckner and McClintock, 1985; Johnston, 1984). These studies found evidence in support of this hypothesis. The schema theory also included scaffolded learning to facilitate the allocation of attention and orderly search in memory to enable elaboration and inferences which permitted reconstruction.

Along with schema theory, the concept of metacognition served to introduce the idea of the strategic reader. When readers encounter comprehension problems, metacomprehension provides them with tools (i.e. metacognitive strategies) to solve these problems. According to the idea of the strategic reader, the reader would select

from a repertoire of strategies those strategies that would enable him or her to achieve comprehension. Metacognition was seen as an extension of schema theory and text analysis. While schema theory and text analysis made use of factual information (i.e. declarative knowledge), the concept of metacognition focusses on procedures (i.e. procedural knowledge) used to process the information. Early work in metacognition focused on two areas of research: one area studied the teaching of explicit strategies for remembering (metamemory); the other area studied the strategies that readers use to monitor, evaluate, and repair their comprehension. Later studies contributed to this field of study by including the concept of conditional knowledge (Paris, Lipson and Wixson, 1983).

After the mid-1990s, reading was studied jointly with writing and listening and further important contributions were made to the field of schema theory and metacognition. McNamara, Miller and Bransford (1991) added to schema theory by proposing mental models. Mental models provide readers with an alternative route or interpretation of the text for uncommon situations (Pearson, 2008). In other words, a mental model results from the interaction between text base (ideas) and prior knowledge. Another theory that set out to improve schema theory was the *flexibility theory* elaborated by Spiro et al. (1987). Similar to McNamara's work, Spiro and colleagues suggested that a more open and flexible approach might help learners. Students should be trained to accept different perspectives on the text and learn to associate ideas instead of accepting a predetermined correct idea.

Further research on cognitive development identified the importance of using context to help students focus their attention in comprehension. Teachers designed activities to help students concentrate on specific information. It was observed that fluent readers, highly skilled in word recognition, do not make use of external information.

In brief, with the development of cognitive psychology, studies started to focus either on describing reading processes or on carrying out experimental work to apply the work of schema theory and metacognitive development. The most important findings of this period (Pearson and Gallagher, 1983) were that a) the conscious focus on text structure benefits students' comprehension, b) relationships created between text content and students' background knowledge permit students to draw inferences and make predictions, c) comprehension improves if reading monitoring is performed, and d) that explicit teaching of reading strategies allows students to understand the text and apply these strategies to problems in the future. The explicit teaching of strategies suggested that the teacher should gradually release

responsibility (Pearson and Gallagher, 1983) to students (i.e. scaffolding).

Comprehension instruction has not only been influenced by the development of cognitive theories but also by social perspectives. Influenced by the social nature of learning, comprehension instruction has been studied in combination or embedded within various other areas such as cultural, political and social sciences. Studies since the early 1980s have led scholars to agree on two main results: a) explicit teaching of strategies can help to improve students' comprehension, and b) the use of strategies often enhances comprehension of new texts, i.e. strategies might be considered as transferable skills.

3.3.1.2. Review of applied studies in reading comprehension instruction

Due to the large number of original studies, this section draws primarily on existing reviews of studies and highlights original studies which greatly influenced the research agenda.

In 1981, Levin and Pressley published the first comprehensive review of studies carried out after Durkin's results. Their analysis showed that research followed two main tendencies: a) studies testing the effect of time at which the strategy was used (before, during or after reading) and b) studies where a person other than the student was applying the strategies (e.g. teachers or researchers). The studies were mainly interventions where students were expected to respond to stimuli controlled by the teacher or the researcher. A drawback of these studies is that they were not designed following real instructional research. This lack of truly experimental research was emphasised by Tierney and J. W. Cunningham (1980), who suggested that the theoretical studies carried out up to 1980 needed to be applied to experimental research in the classroom; i.e. in studies which could address the complexity of context, teaching and classroom.

In a later update, Tierney and J. W. Cunningham (1984) adopted a similar reviewing approach to the one followed by Levin and Pressley. They reviewed instructional procedures in the different reading phases (pre, during, and post), but they also differentiated the aim of the procedures. They suggested that research could be categorised into two groups: a) studies that examine ways to increase comprehension from text and b) studies that examine ways to increase student's ability to comprehend and learn. The former included instructional approaches mainly guided by the teacher on a specific passage as well as activities that could be implemented at any reading stage of the text. The latter aimed at enhancing comprehension strategies.

Some of the most reviewed strategies in Tierney and J. W. Cunningham (1984) paper were: engaging prior knowledge, identifying task demands, summarising, and training in metacomprehension.

Although Tierney and J. W. Cunningham (1984) reported encouraging results from the studies, they were reluctant to accept the usefulness of direct teaching. They pointed out the narrow focus on these strategies and the apparent mechanical approach. In spite of a wide variety of studies included in the review, most of them tended to focus on texts where prose was the main source of reading material. This might limit the transferability of the findings to other settings where prose is not the main text type. The authors' recommendation for future research included the integration of readers' needs and characteristics as well as those of the learning groups and the teachers into the design of future studies. They also recommended the inclusion of teachers as collaborators in research. This would bring classroom reality into research, thereby helping to close the perceived gap between teaching practice and educational research.

Pearson and Gallagher (1983) reviewed and evaluated studies related to instructional practises carried out between 1978 and 1982. This review classified the studies based on their methodology: a) studies that described the current situation of instruction in schools or curricula, b) studies which proved existence of relationships among variables, with special focus on different types of readers' abilities and ages, c) studies that increased students' ability to comprehend a given text, and d) studies that evaluated the influence of a variable when used as part of a curriculum. An important feature of this review is that Pearson and Gallagher support the 'model of explicit instruction'. In this model, the teacher demonstrates how to carry out a strategy and engages the students in guided practice followed by independent practice. In the late stages, students apply the strategies learnt on their own while reading regular texts. This model was an adaptation to Campione's model of 'gradual release of responsibility' (Campione, 1981; Pearson and Gallagher, 1983). Another important finding was that the majority of studies did not success in establishing causality between institutional practice and outcomes.

In 1983, Paris, Lipson and Wixson added to the field the label of metacognitive instructional research, helping to differentiate between instruction and instructional research. They are regarded as the first authors to speak of conditional knowledge and to outline the differences between declarative and conditional knowledge, which J. R. Anderson (2009) had previously described.

The reviews mentioned above evaluated scientific research carried out by academ-

ics. Other reviews looked at studies that have helped to disseminate the knowledge amongst practitioners. Pearson (1985), for instance, released a review to explain to reading educators how to use the new research findings as a basis for curriculum change in schools. Similarly, (Gersten and Carnine, 1986) review of the literature was targeted at administrators and curriculum specialists. Even policy makers, with reports such as 'Becoming a Nation of Readers' (R. C. Anderson et al., 1985) attempted to explain the instructional implication from reading research by emphasising the role of comprehension instruction. One of the key papers was that of Pressley, Goodchild et al. in 1989, which focussed on forty major pieces of strategy instructional research evaluating them for methodological adequacy of internal and external validity, assessing and transference

Pearson and Fielding (1991) focused their review on studies designed to determine methods of comprehension instruction associated with improvement on some measure of comprehension. These authors distinguished studies that improve students' comprehension of texts from those that improve students' ability to comprehend texts independently. Unlike other reviews, Pearson and Fielding divided the studies according to the kind of texts, text features, and strategies. This new classification included studies not previously analysed in the major reviews. The first part included studies working with comprehension improvement of narrative and expository texts. The second part reviewed some generic strategies and practices derived from the studies in the first part such as self-questioning, self-monitoring and peer interaction.

The work carried out until the 1990s focussed on comprehension instruction as well as on related areas such as strategies used by good and poor readers, comprehension strategies, and comprehension development amongst others. However, it was recognised that *"[...] instructional researchers focused on the aspects of comprehension most relevant to the comprehension instruction they favored [...]"* (Pressley, 2000, p. 545). That is, the topics were often selectively chosen and discussed in isolation, which might be considered as a drawback of the research carried out at that time. This problem was identified by Pressley (2000) who attempted to bring together the knowledge of all these areas. He stated that appropriate instructional recommendations would result from the understanding of effective comprehension development.

Based on the analysis of comprehension development, Pressley (2000) suggested instruction in three main areas to help less skilled readers to increase their comprehension: improving word-level competencies, building background knowledge, and promoting the use of comprehension strategies. However, instruction in the

first two areas may be hampered by the relatively young age of learners and the consequent lack of prior knowledge and schemata. Development of decoding skills is generally promoted in the early years of schooling. The schema theory suggests that readers rely on their schemas (i.e. prior knowledge) to draw inferences (e.g Pearson, 2008). Hence, the more schemata a reader possesses, the easier it is for him or her to resolve comprehension problems. However, background knowledge is acquired over the course of a person's life by means of enriching experiences and exposure to the world; thus, building up rich background knowledge requires time and contact with meaningful experiences. Given these limitations, it is perhaps not surprising that studies looking to increase comprehension through vocabulary instruction have not always yielded positive results (Pressley, 2000).

While word decoding and prior knowledge are regarded as automatic and unconscious processes , strategies are regarded as being controlled by readers. The use of comprehension strategies has been well researched, with a view to find ways to teach less skilled readers how to consciously make use of reading strategies to improve their comprehension. For instance, Pressley and Afflerbach (1995) showed that skilled readers make use of some strategies while reading texts. They analysed verbal protocols of mature readers (40+) while reading a text and listed some of the most used reported strategies. These include:

- setting a purpose for reading,
- overviewing the text to evaluate its relevance,
- reading selectively,
- associating the text with their prior knowledge,
- evaluating hypotheses set during the initial overview,
- solving inconsistencies in the text based on their previous knowledge,
- creating meaning for new words relevant to the meaning of the text,
- rereading,
- underlining,
- making notes,
- paraphrasing,

- evaluating the text,
- reviewing after finishing,
- and finding a use for the information in the future.

The first part of the literature review has drawn attention to the importance of comprehension and to the efforts made in the last decades in order to understand such a complex process. Following the suggestion made by Pressley, the second part reviews the research done in the area of reading comprehension strategies. Instruction in comprehension strategies may improve the abilities of less skilled readers, an aim central to the intervention developed in this study.

3.3.2. Reading comprehension strategies

This section defines the term strategies and investigates the characteristics of reading comprehension strategies. First, it provides an overview of the research that has been carried out in the area aiming to enhance comprehension. Then, it reviews the different approaches developed to teach reading comprehension strategies. Finally, it identifies the gap in previous reading comprehension research and thereby places this study in the context of the existing literature.

3.3.2.1. Strategies

Scholars interested in cognitivism defined all the cognitive processes as 'strategies' (Pressley, Goodchild et al., 1989). Over time the term strategy was spread and it is now widely misused or misunderstood. One of the main problems is that the term is used as synonymous with skills. Although skills are also cognitive processes, these are considered as *"automatic actions [..][which are performed] with speed, efficiency, and fluency"* (Afflerbach, Pearson and Paris, 2008, p. 368). The term reading strategy has been redefined and adjusted to fulfill authors' needs. Several definitions share the principle that reading strategies are cognitive operations that help *"to enhance reading comprehension failures"* (Singhal, 2001, p. 2) - see also (Hacker, 1998). Unlike skills, strategies are conscious and controllable processes used to 'work towards a goal' (Hacker, 1998; Grabe and Stoller, 2002; R. Pritchard, 1990; Afflerbach, Pearson and Paris, 2008). Strategies can be used in different contexts so they can be used and adapted to a specific situation with positive results (Afflerbach, Pearson and Paris, 2008). It is worth noting that, practice might lead strategies to become skills in the

same way that a skill used in a conscious way may serve as a strategy (Afflerbach, Pearson and Paris, 2008). In brief, the main difference between skills and strategies lies in the reader's awareness and control of the processes. Reading strategies are perceived as problem solving tools that help readers to overcome comprehension problems. Whenever the word strategy appears in this work, the author refers to the definition of Harris and colleagues:

> *A strategy is a special form of procedural knowledge that is intentionally purposefully and effortfully applied to a given task or situation for which one's typical or automatic pattern of thought or behaviour is perceived as inadequate or nonoptimal* (Harris, Alexander and S. Graham, 2008, p. 89).

Reading strategies are used by readers as self-regulated tools which allow them to have control over information and their own learning (Schunk, 2000). Strategies have been classified into cognitive and metacognitive strategies. Cognitive strategies are perceived as processes that facilitate the interaction of the reader with world knowledge in order to construct meaning (Akyel and Erçetin, 2009; Hacker, 1998; Pressley, 2000). Metacognitive strategies are processes which allow to control and monitor this interaction (Hacker, 1998). The following section provides a summary of how reading strategy research has evolved. It builds on research trends that emphasise the characteristics of the strategies and reviews what has been learnt during as part of these efforts, and identifies gaps in the existing literature.

3.3.2.2. Trends of research in strategy instruction

Strategy instruction is a widely researched area with several studies sharing similar characteristics. Strategy instruction research has been divided into three periods (1980s, 1990s, and current) (Raphael et al., 2008; Pressley, 2000). Following this approach, the next section analyses the important topics in the field, shows the development of the area, and the gaps left for future research.

After the observation made by Durkin (1978) regarding the lack of comprehension instruction in schools, research followed two paths: a) studies focussing on the strategies used by skilled and non-skilled readers and b) studies searching for ways to teach effectively the strategies identified in the first path. One of the defining characteristics of most studies in this period is the implementation of isolated strategies; that is, teaching one strategy at the time. These studies normally included two experimental groups: a treating group receiving an intense training and exposure to the

strategy, and a control group exposed to the same texts and assessment but following the normal methodology. Researchers measured the effectiveness of the intervention, carried out outside the classroom, by designing their own outcome measures, such as retelling a story or multiple-choice items. Most of the studies showed comprehension improvement for the students who had been taught the strategy. The general agreement was that readers in control of their strategies had better comprehension; hence strategy instruction was deemed necessary to enhance comprehension of less skilled readers (Raphael et al., 2008). Some of the individual strategies considered as effective were prior knowledge activation, question generation, construction of mental images, predicting, identifying important information, summarising, making inferences, analysis of stories, among others (see the table of strategies and studies in Appendix B.1).

Some of the studies looked to improve comprehension of the text used in the intervention, while others aimed to find ways to improve students' ability to understand different texts (Tierney and J. W. Cunningham, 1984). In the interventions, researchers normally used declarative and procedural knowledge. Declarative knowledge or knowing the 'that' (Schunk, 2000) was defined by telling students that a strategy was a tool which helped them to improve their comprehension. Procedural knowledge or knowing the 'how' was used by teaching students 'how' to use the strategy. It referred to the steps to follow in order to perform the strategy. Studies generally found that, despite students having the knowledge, they were unable to decide the appropriate moment to use it. This knowledge was always provided by the teacher or the researcher who had control over the rules and the procedures to apply the strategy.

Paris, Lipson and Wixson (1983) introduced the concept of conditional knowledge. In this study the authors taught participants to identify the appropriate situation to apply the strategy. Conditional knowledge provided students with a rationale for using the strategies and allowed them to have greater control of the strategies by knowing 'when' and 'why' to apply the taught strategies. Conditional knowledge provides students with independence because they can decide when a strategy is useful without depending on the teacher; that is, students manipulate the declarative and procedural knowledge (Schunk, 2000). Also, students become aware of their needs and are able to choose among a variety of options (i.e. strategies). Although Paris, Lipson and Wixson (1983) introduced the idea of conditional knowledge, it was not developed and studied until later.

Another concern at the time was the need to find an effective way to teach

students to use and control their knowledge. Scaffolding (Pearson, 1985) addressed this problem by providing students with gradual increase of responsibility. During scaffolding learning the teacher explicitly introduces a concept or strategy and explains how it works and why it is important. The instruction is normally provided by using think alouds and modeling by the teacher. After this introduction, the students have the opportunity to try the strategy by themselves in small groups. This support continues until the students are able to perform the strategy independently. This same model has been represented in the Gradual Release of Responsibility Model (Au and Raphael, 1998; Raphael et al., 2008).

The results from the studies testing isolated strategies and the ideas of Paris et al and Pearson were combined in further studies to design complex interventions to test groups of strategies. In addition, instruction in these new studies was carried out within classroom settings. Some of these studies included a social aspect. A combination of theories was the basis of several approaches which shared common characteristics such as *"modeling, scaffolding, guided practice, and independent practice of the strategies all within the socially constructed nature of classroom discourse"* (Raphael et al., 2008, p. 453). Some of the most important approaches such as reciprocal teaching, transactional strategy instruction, etc are briefly described below.

3.3.2.3. Teaching approaches to reading comprehension

3.3.2.3.1. Collaborative Strategic Reading

The Collaborative Strategic Reading (CSR) approach teaches students how to become strategic readers. It provides students with cognitive instruction to understand texts in content areas. It also assists students with learning disabilities, and provides them with opportunities to work in collaborative environments. Four strategies are normally taught to allow internalisation and memorisation of the strategies. After the strategies have been taught, students work in small mixed-ability groups to practise them. A key element of the approach is the work in small collaborative groups (Raphael et al., 2008).

3.3.2.3.2. Peer-Assisted Learning Strategies

The Peer-Assisted Learning Strategies (PALS) approach was designed to enhance reading fluency and reading comprehension of low-achieving students with and without disabilities. It is based on a peer tutoring model to teach students to

summarise, retell, monitor, elaborate, and predict. After direct teaching of the strategies, the responsibility of using the strategies is gradually passed to students, who work in pairs with one of them acting as a coach. The use of prompts and feedback is used to provide support to students (D. Fuchs and L. Fuchs, 1997; Mathes et al., 1998).

3.3.2.3.3. Transactional Strategies Instruction

Transactional Strategies Instruction (TSI), developed by Pressley, El-Dinary et al. (1992), emphasised the importance of direct explanation. This instructional approach was designed to take reading strategy research to schools; hence, it was delivered across the curriculum. It involves direct explanation of reading strategies for long periods of time; typically one or two years. The instruction consists of modeling of strategies by the teacher, as well as giving explanations using the conditional knowledge to explain the value of strategies to students. An important characteristic of this model is the transference of strategies through practice. After learning a new strategy, students have the opportunity to apply it to another text in another context. Teaching a repertoire of strategies and the development of metacognition contribute to self-regulated cognition and an increment in the word knowledge as well as student motivation (Baker and Carter, 2008). Unlike previous approaches, students are responsible of their needs and the teacher helps them to make predictions and interpretations. Strategies are taught one at the time but they are all used and revised throughout the instruction. Work is carried out in student-centered reading groups where students discuss the tasks and share their experiences to become aware of the different ways of thinking. The role of the teacher is to coach by providing hints and clues to help students elaborate and discover new ways to solve their comprehension problems. However, two points are considered as drawbacks of the approach: a) no explanation is given about the benefits of the strategies and b) the length of the intervention. The authors point out the practical problems of carrying out and evaluating a long intervention, e.g. time taken from the curriculum, attrition rates and the effect of time on outcomes unrelated to the intervention (maturation). In addition, ethical issues about delivering long experimental interventions have been raised because students are either exposed to an ineffective new teaching approach, or an effective approach is withheld from students in the control group over a prolonged period of time (Pressley, 2000).

3.3.2.3.4. Concept-Oriented Reading Instruction

Concept-Oriented Reading Instruction (CORI) was designed to engage intrinsically motivated students proficient in the use of strategies to build knowledge through a variety of texts. It combines strategy instruction with motivational features and the opportunity for collaboration. The strategies taught include activation of background knowledge, question elaboration, searching for information in multiple texts, summarising, and organising information graphically. So far, it has been delivered mainly in primary education settings.

3.3.2.3.5. Reciprocal teaching

The reciprocal teaching approach builts on Vygotzky's zone of proximal development theory. In this approach the child first acts as an apprentice and, thanks to experience and interaction with a more experienced interlocutor, subsequently begins to perform more complex tasks. Reciprocal teaching was defined as *"a set of learning conditions in which children first experience a particular set of cognitive activities in the presence of experts, and only gradually come to perform these functions by themselves"* (Brown & Palincsar, 1989 in Rosenshine & Meister, 1994, p.480). Students read a passage of expository material paragraph by paragraph; during the reading they learn and practice four strategies and predict what might appear in the next paragraph. In early stages, the teacher assumes the responsibility for instruction by explicitly modelling the process. While students practise the strategies, the teacher supports each student through specific feedback, additional modelling, coaching, hints, and explanations. During the guided practice, reciprocal teaching makes use of the dialogue in an attempt to gain meaning from the text. Discussions serve to elaborate or comment on another student's summary or predictions, suggest other questions, request clarification and help to solve misunderstandings. The teacher acts as a guide providing support when needed. There is a gradual shift of responsibility with emphasis in cooperative work. In addition, students are provided with instruction in why, when, and where such activities should be applied when facing a new text. In sum, the main features of this instructional approach are the instruction and practice of four comprehension strategies and the use of reciprocal teaching dialogue as a vehicle for learning and practising the strategies. To date, reciprocal teaching is one of the most studied approaches and its procedures have been widely replicated. What follows is a summary of a systematic review of studies applying this approach.

Rosenshine and Meister (1994) reviewed sixteen studies based on reciprocal

teaching - including published and unpublished quantitative studies indexed in Dissertation Abstracts International. The authors looked for studies using expository materials which contained experimental and control groups. The number of strategies taught varied across studies. Whereas twelve studies taught four strategies, the remaining four studies taught between two to ten studies. Only four of the studies were peer-reviewed and published, the rest were dissertations or papers presented at conferences. The review describes two types of reciprocal teaching: reciprocal teaching only (RTO) and explicit teaching before reciprocal teaching (ET-RT). The difference is that in the original study (RTO), the modelling and instruction take place during the dialogues. The quality of the studies was defined by looking at the design of the study and when possible the quality of the study dialogues. The studies were classified as high, medium or low quality. Classroom teachers were trained by the researchers in seven of the studies, in the rest the researcher delivered the instruction. A drawback of the studies is that no criterion was set to evaluate the quality of the instruction. Moreover, lack of observation of instruction was a common problem. Results from the studies using standardised tests as outcome measures were significant in only two of nine studies. In contrast, six of the seven studies employing experimenter-developed tests had significant results. Overall, the effect size was higher (.88) when experimenter-developed comprehension tests were used, and lower (.32) on standardised tests. No relationship was found between the number of strategies taught and student achievement. Studies were equally effective regardless of the person providing the instruction (e.g. researcher or teacher). In eight of the studies the researchers assessed the extent to which the participants had learned the specific strategies. The greatest effect was found when the strategies had been explicitly taught before the instructional approach and in the studies with experimenter-developed tests. Rosenshine and Meister (1994) concluded that not enough information had been provided about the implementation of the approach in the research reports.

3.3.2.3.6. Cognitive Academic Language Learning Approach

The Cognitive Academic Language Learning Approach (CALLA) developed by Chamot and O'Malley (1987) is based on the repetition of a five-step cycle. The cycle consists of introduction, instruction, practise, evaluation and application of the chosen strategies. The introduction allows students to set objectives for reading and to identify the cognitive strategies they already use. Students identify the value

of these strategies and are introduced and guided to the use of the metacognitive strategies. In the second stage, the teacher introduces the strategies by modelling the procedures and explicit instruction of how to use them is provided. During the practice, the students have the opportunity to try the strategy on authentic materials. The teacher monitors the practice and provides assistance when needed by providing students with clues. Teachers' support helps students to recognise when a strategy has not given the expected results. The evaluation provides opportunities for students to assess their success in the task. Through self-questioning and discussions, students become aware of how the effectiveness of strategies differs depending on the type of activity carried out. In the last stage, students are encouraged to summarise their new knowledge and to try out the strategies in new contexts as well as to explore the combination of strategies according to their own preferences.

3.3.2.3.7. Summary of teaching approaches to reading comprehension

The discussed teaching approaches have all showed significant results in a variety of settings. Although individual drawbacks have been discussed alongside the descriptions, a common problem was identified: strategies used in the instructional approaches are imposed onto students. This might prevent students to become responsible for their knowledge. Paris, Lipson and Wixson (1983) suggested that for strategies to be effective, students should become aware of the functionality of the strategies. Furthermore, students need to have a repertoire of strategies to choose from the appropriate one for the presented task. Another element preventing comprehension is the small number of strategies taught in some of the interventions. This might limit students' options to choose the most effective strategy for a defined task.

There are also several common characteristics that may have contributed to the observed improvements in students' comprehension. These characteristics are listed below because they will be included in the design of the present study:

1. *Teacher explanations and mental modeling* - The teacher explains what the strategy is and how it works, followed by an example within a real and clear context. He models the process by thinking aloud all the processes that he follows to take a decision and the rationale behind it. This procedure helps students become aware of how the silent processes would work when the learners are working independently.

2. *Practice in real contexts* - Because comprehension can be developed, practice

provides students with opportunities to try whether a strategy is effective or not while learning and developing their own theories about every strategy. Not all texts can be addressed with the same strategy, thus the use of real contexts provides a variety of situations where students practise and test the strategies in the most possible real situation.

3. *Monitor and additional explanations/modeling* - After trying out the strategies, students might have questions that need clarification. The teacher provides explanations and gives further examples based on different texts to reinforce the understanding of the learned strategies

4. *Feedback* - In order to allow development and improvement, constructive feedback is needed. This feedback must always look for encouragement to emphasise students' motivation and help them see the importance and effectiveness of the strategy. The amount of feedback is reduced once the students begin to use the strategies in an appropriate way.

3.3.3. Review of recent studies that are closely related to the current study

Studies in the last decade have been conducted in real classroom settings to enhance the ecological validity of the findings. Stronger emphasis was given to the teaching of metacognitive strategies and the use of self-regulation. Most of the studies have been designed to measure comprehension when the instruction is given by the classroom teacher. The following section reviews in detail seven recent studies that have implemented very similar research designs and sought to answer similar research questions as the current study. These studies were highly influential for the design of the current study and served, for example, as a starting point for the selection of strategies to be taught (see Appendix B.1 and Section 4.6.2).

Boulware-Gooden et al. (2007) conducted a five-week intervention to test the effectiveness of systematic direct instruction of metacognitive strategies on comprehension and vocabulary development. Using the eight most effective strategies suggested by the National Reading Panel in 2000, direct instruction was given to 119 third-grade students of two urban elementary schools in the United States. The lessons were divided into five parts: introduction, vocabulary teaching, story reading, summary and questions. Direct instruction of the strategies was delivered in sessions of 30 minutes for 25 days. The control group was not encouraged to think aloud while reading and did not identify elements of the expository passage. Results from a

standardised comprehension test suggest that the use of metacognitive instruction improved students' academic achievement.

Cubukcu (2008) worked with 130 third-year teacher trainees. Intact groups were treated under experimental and control conditions. The experimental group received 45 minutes of instruction on 10 metacognitive strategies for a period of five weeks. The instruction followed the steps of the Cognitive Academic Language Learning Approach (CALLA). Results from a standardised test showed that the systematic explicit instruction of metacognitive strategies improved the comprehension of students in the experimental group.

Houtveen and Grift (2007) carried out an experiment where teachers of ten-year-old children were trained to teach metacognitive strategies. The teachers in the experimental group were trained in metacognitive strategy instruction and in optimising instruction time for reading comprehension. Teachers in the control group did not receive any training. The duration of the intervention is not clearly stated.

The effectiveness of the intervention was measured with respect to gains in the use of metacognitive strategies. In addition, an observation instrument was used to measure teachers' behaviour during the instruction and a standardised test served to measure students' reading comprehension at a delayed follow-up after one year. Students in the experimental group outperformed their peers in the control group on use of metacognitive strategies. The follow-up measure in the next academic year showed a difference in terms of reading comprehension in favour of the students who had been taught by teachers in the experimental group. This may indicate that the effect of the intervention is maintained over time, although the follow-up assessment focused on a different outcome (comprehension) than the immediate post-test (strategy use). The result of the observation instrument suggests that teachers in the experimental group demonstrated better metacognitive strategy instruction.

Inchausti de Jou and Sperb (2009) conducted an intervention in an elementary school in Brazil. An experimental group and three control groups of fifth-grade students were formed. Students in the experimental group were instructed about superstructure and macrostructure of the texts as well as cognitive and metacognitive strategies. Control group 1 read the same materials as the experimental group but without receiving instruction (active control group). Control groups 2 and 3 did not receive any modification to its normal classes (passive control group). Students from the third group belonged to a different school and were included in the study to control for school effects. Students' comprehension was measured through memory tasks: recall and multiple-choice questions. The students worked with narrative texts

in twelve 45-minute sessions, divided into two parts. The strategies were explained in the first part and group practice was allowed during the second part. Results did not show significant differences among the groups in the use of cognitive strategies. However a significant difference was found for the experimental group in the use of metacognitive strategies. There was a statistically significant difference in favour of the experimental group when reading comprehension was assessed using the recall approach; however, no difference was found in the multiple-choice answers. The authors concluded that students in the experimental group developed their declarative and procedural knowledge of macrostructure and superstructure as well as their comprehension.

Thomas and Barksdale-Ladd (2000) delivered intensive instruction on how to improve metacognition to ten undergraduate students in their fourth year of studies. These undergraduates instructed young pupils at local schools. Think-aloud analyses and discussions added to the instruction were intended to help students to document their understanding of the metacognitive strategies. Furthermore, the participants wrote journals about their experience in teaching younger students. These journals were analysed to document the application of strategies when tutoring children and see whether the participants had passed their knowledge about metacognition on to their students. Results showed that the intervention had a powerful impact on the development of metacognitive strategies in the undergraduate students. However, the authors found no evidence that this knowledge was passed on to their tutees.

Wilawan (2007) instructed 60 Thai undergraduate students learning English on identifying the main idea of a given text. The intervention lasted 15 hours and was based on procedures of reciprocal teaching. The EFL students were divided into three treatment groups to see whether the instructional procedure affected their main idea comprehension. The first group received training on lexical cohesion and metacognitive strategy training, whereas the second group received only metacognitive strategy training and the third group worked with traditional skill-based instruction. Results indicated that although students from the three groups improved their performance on main idea comprehension test, only gains for the groups receiving metacognitive strategy training were statistically significant. Results from the verbal protocols revealed differences in pattern strategy use between the two groups who received instruction. Students receiving training on lexical cohesion reported using more bottom-up strategies; whereas the group receiving training in metacognition used more the top-down strategies.

The review of these studies allows identification of two distinct classes of studies:

first, those that are carried out in elementary schools where English is the native language; second, studies that are conducted in secondary or higher education where English is taught as a foreign or second language. The common characteristic of both types of studies is that they focus entirely on the comprehension of texts in English language and it is not clear in how far results can be generalised to other languages.

Only one study was identified reporting the results of an intervention for native Spanish speakers (Madariaga Orbea and Martínez Villabeitia, 2010). The study was conducted in the town of Vizcaya, Northern Spain. The researchers developed a program where the teachers were trained to directly teach reading comprehension and metacognitive strategies, both in Spanish and Euskera, the local dialect. The program was divided into two phases: the training of the teaching staff and the program application. Teachers were trained in three 60-minute meetings. They were taught the need of teaching to understand, strategies implied in the effective development of reading comprehension, theoretical and practical grounds of direct instruction and guided reading of the direct instruction program. The lessons consisted of 5 stages: introduction, example, direct instruction, application steered by the teacher and practice. Third and fourth-grade students participated in the program. The participants taught in Spanish were allocated into two groups. The experimental group worked with strategies addressing the construction of text macrostructure, structural strategy and metacognitive strategies. The control group followed the traditional instruction. The same procedure was followed for the other two groups who received the instruction in Euskera. Students' vocabulary was tested before the intervention. The instrument showed a slight difference in vocabulary knowledge between the two groups. Students' comprehension was measured using a standardised test. Results from the pretest showed an initial difference between groups in the levels of reading comprehension. Post-test results from students scoring between the 75th and the 25th percentile were analysed. Results showed that both groups had improved their average reading comprehension. The experimental group made a significant improvement. An Analysis of Covariance (ANCOVA) was carried out to control the influence of personal and context variables. The pretest results were included as a co-variant. Results of the multivariate analysis indicated the level of Spanish vocabulary as a significant factor affecting the results of the post-test. For the control group, pupils' attitudes towards reading also influenced the results. Based on the results of the ANCOVA, the authors concluded that the program was effective in improving students' reading comprehension. Although the study shows improvement in comprehension, the effect size is not reported. Therefore, it is difficult to judge

how effective the intervention was i.e. the magnitude of the effect. Furthermore, it does not specify whether the topics of the texts used were familiar or unfamiliar to students.

The analysis of the studies reviewed in the previous sections provides evidence about the effectiveness of the reading strategies. Research suggests that comprehension can be developed through direct instruction of reading strategies (Carrell, 1989; Pressley, 2000). Therefore, it might be hypothesised that the direct teaching of reading strategies might also foster the comprehension of native Spanish-speaker students; specifically, low-skilled readers struggling to understand texts written in Spanish. Madariaga Orbea and Martínez Villabeitia (2010) study was the only study identified that has investigated a related hypothesis. The current study extended this work to participants of a different age using different texts (i.e. expository texts). The principal objective of the study is to assess whether the results of previous studies carry over to the Mexican setting. In order to do so, we conduct a teaching intervention with direct teaching of cognitive and metacognitive strategies to improve reading comprehension in 15-year-old Mexican students.

Several factors were taken into consideration before selecting 15-year-old students to conduct the intervention. The fist and most important one is the low performance of Mexican students in national and international evaluations (Chapter 2). Another was the lack of reading strategies instruction in the national syllabus for secondary schools. The last one is the accessibility to the school and the desire to help students from poor socio-economic background. Students in their last year of secondary school take a national test at the end of the year. Results from this test serve to determine access to higher education colleges. The intervention might help the participants to have a better performance in the selection test, which will result in their access to better colleges.

4. Methodology

4.1. Aim of the study

The evidence from the studies discussed in the literature review (Chapter 3) suggests that the students' comprehension can be improved by the direct teaching of reading strategies. In addition, making students aware of their thinking while reading and teaching them to monitor their understanding helps to improve their comprehension. However, the vast majority of these studies focus only on comprehending texts written in the English language, taught to both native and non-native speakers. It might be hypothesised that the direct teaching of reading strategies also fosters the comprehension of Spanish texts among Spanish speakers, although the empirical evidence remains scarce (Madariaga Orbea and Martínez Villabeitia, 2010).

The aim of this study was to test whether the use of reading strategies enhances comprehension in secondary students in a Mexican public school. This study followed the approach taken by previous research by applying the direct teaching of a set of reading strategies. However, the context was different in that the intervention was applied to native Spanish speakers in a developing country, and reading strategies were taught (and applied) in Spanish. The intervention included direct instruction, delivered first by the researcher and then by the regular class teachers, in several cognitive and metacognitive strategies embedded in a collaborative environment.

This chapter will a) present the research questions, and the rationale behind the methods used to address the questions, b) describe the research approach adopted, including the sampling methods, procedures for data collection and the instruments designed to measure the outcomes, c) explain the teaching procedures and the rationale behind the selection of the reading strategies taught in the intervention, d) outline the procedures used for the data analysis, and e) describe the pilot study that was carried out before the intervention.

4.2. Research questions

The comprehension problems faced by Mexican students are outlined in Chapter 1 and Chapter 2 of this thesis. The PISA results clearly demonstrate the comparably low level of students' reading ability in Mexico compared to many other OECD countries. From the review of the literature, it is evident that the direct instruction of reading strategies enhances comprehension in English students. By training 15-year-old Mexican students in the use of reading strategies, this study aims to answer the following research questions:

- Q1: Does the direct teaching of cognitive and metacognitive reading comprehension strategies help 15-year-old Mexican students to improve their reading comprehension of Spanish texts?
- Q2: Does the intervention raise the students' self-reported awareness of reading strategies and the reported frequency of use?
- Q3: Does the intervention increase the reported 'online' use of reading comprehension strategies when reading unfamiliar texts?
- Q4: Is there a relationship between the self-reported use of strategies and the students' performance on the reading comprehension test?
- Q5: Do other factors, such as indicators of socio-economic status, baseline reading comprehension scores and self-reported use, as well as word recognition, influence the effectiveness of the intervention on reading comprehension and strategy use?

4.3. Research framework: mixed methods

Two different research methodologies could be used to answer the research questions: quantitative and qualitative analysis. The former combines the observation and measurement of quantitative data with statistical analysis to isolate the effect of an intervention on an observed outcome. To this end, surveys, experiments and tests are used to verify theories or explanations. The latter relies on case studies, grounded theory, narrative and open-ended questions to gain a more in-depth understanding of the processes in context.

There has been a constant and on-going debate about the strengths and weaknesses of both approaches (e.g Howe, 1988; Nathaniel, 2007; Sale, Lohfeld and Kevin,

2002). However, it is increasingly being accepted that, by combining both methods within the same study (the 'mixed methods approach'), researchers can increase the strength of their study and allow for better inferences and a more complete analysis of data (Tashakkori and Teddlie, 2003). Tashakkori and Teddlie define mixed methods research as *"studies [using] qualitative and quantitative data collection and analysis techniques in either parallel or sequential phases"* (Tashakkori and Teddlie, 2003, p. 11). Although a mixed methods approach has often implicitly been employed throughout the history of social science research, it was not until recently that a group of researchers provided clear definitions and developed methodologies for it, allowing a more rigorous use of this approach (Creswell, 2003; Morse, 2003; Tashakkori and Teddlie, 1998; Tashakkori and Teddlie, 2003).

Following the discussion in the literature, and given the perceived advantages of the mixed methods approach, it was decided to implement this approach in this study. Quantitative data allowed the testing of word knowledge, the students' comprehension and the awareness of strategy use. Qualitative data, collected through verbal reports, provided a closer observation of the students' behaviour. The extent to which the different types of data converge provided additional evidence about the usefulness of direct reading comprehension instruction in 15-year-old Mexican students.

Creswell (2003) proposes six research strategies as a guide to researchers in designing a research model. The current study followed the 'concurrent triangulation strategy', using quantitative and qualitative methods separately initially, then combined in the later stages. As such, the concurrent triangulation strategy allows researchers to *"confirm, cross-validate or corroborate findings"* (Creswell, 2003). Both the qualitative and quantitative data collection in this study were conducted in parallel. Data were collected during three phases: before the intervention (pre-test), immediately after the intervention (immediate post-test) and four months after the intervention (delayed post-test). Both quantitative and qualitative methods were used, though the majority of the data collection and analysis were quantitative. The data were analysed separately and then the results were integrated and interpreted jointly.

Figure 4.1 outlines the design of the study by presenting the instruments used during the data collection alongside the analyses carried out to answer the research questions. It is worth noting that data labelled 'Word-level study' (WLS) were collected only once, before the intervention. The aim of this phase was to assess the students' knowledge of low-level reading skills in order to investigate any negative in-

fluence that this might have on reading comprehension. The reading comprehension tests, metacognitive assessment instrument (MAI) and verbal reports were carried out before and after the intervention. The design of each instrument is explained in detail in Section 4.5.

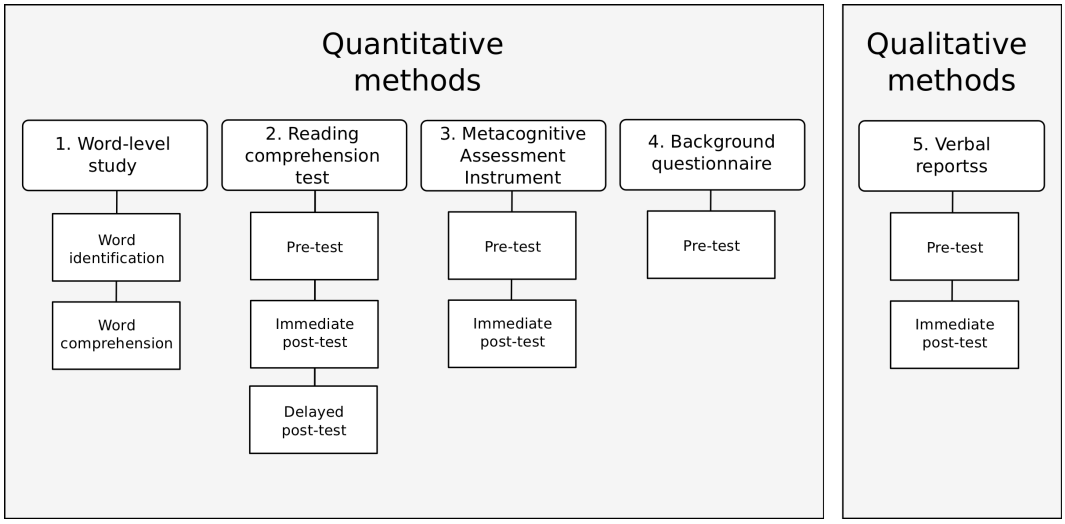


Figure 4.1.: Instruments used for data collection

To answer the first research question (Q1) and test whether the direct teaching of the reading strategies helped to improve the students' comprehension, the students' performance was assessed through a reading comprehension test designed by the researcher, based on items from the PISA tests. The scores of the comprehension tests before, immediately after, and four months after the intervention were compared.

Changes in the students' awareness of strategy use (Q2) were assessed by comparing the metacognitive assessment instrument (MAI) responses before and after the intervention. This instrument is based on a four-point Likert-scale and measures the students' reported strategy use and awareness. A drawback of using a self-reported instrument is that the participants might overestimate or underestimate their strategy use, as they may not understand the questionnaire items in the way intended or may have insufficient self-awareness to report their own behaviour correctly (L. Cohen, Manion and Morrison, 2011). To obtain a more objective measure of the students' strategy use, a subsample of students was asked to perform several reading tasks before and after the intervention and told to think aloud during the tasks (verbal reports). The researcher monitored the students during these tasks and the resulting data were then analysed to answer the third research question (Q3).

To answer the fourth research question (Q4), the results from the comprehension tests were analysed in conjunction with data from the self-report strategy-use questionnaire, and the verbal reports. The gain scores were compared with the self-reported use of strategies to test whether a relationship exists between the use of strategies and comprehension improvement.

Information obtained via a background questionnaire was used to answer research question Q5. It served to identify possible confounding factors that might have influenced the results. These factors were included as covariates in an analysis of reading comprehension scores and MAI responses to explore their effect and potential role as confounding factors.

Sections 4.5 and 4.8 provide further details of the steps followed to design and test the instruments used in the study.

4.4. Research approach

4.4.1. Participants

A secondary school located in a low-income working class neighbourhood in Mexico City was designated by the Mexican Secretariat of Education to carry out the project [1]. The school is enrolled in a project called 'Escuelas de tiempo completo' (full time school). Students enrolled in these schools spend more hours at school than those enrolled in general schools. Core lessons are delivered from 7:00am to 2:30pm and are followed by workshops, such as theatre, music, or crafts. In addition, time is allocated for students to complete their homework.

All students in their last year of compulsory education were selected to participate in the project. The study sample contained 177 students, 89 of whom were female (52.4%). The average age of the students was 14.5 years at the beginning of the study, which is approximately the age at which students take PISA exams. The project started during the second week of the academic year (August 2011). The students were informed that a reading workshop would be added to the evening activities. They were also told that they would be undertaking reading activities to help them to improve their reading skills in order to perform better during the high

[1] To carry out research in schools in Mexico, it is compulsory to ask for permission from the Secretariat of Education. Based on the research proposal for the project, the secretariat designated a school to participate in the project. The selection of the school was therefore not under the control of the researcher.

school's entry test at the end of the year [2]. However, the actual purpose of the study was not revealed to the students to avoid improvement due to sensitization to the experimental conditions (L. Cohen, Manion and Morrison, 2011).

The students formed five intact classes with an average class size of 35 students. The classes had been formed at the beginning of the academic year by the psychologist and the social worker of the school based on the results of the Differential Aptitude Test (DAT) [3]. The DAT is a psychological test that measures individuals' abilities and aptitudes. It consists of eight tests which measure verbal reasoning, numerical ability, abstract and mechanical reasoning, space relations, clerical speed, accuracy and spelling and language usage. Students were allocated in a way that ensured that each class had a mix of students with different levels of ability but that their ability was approximately equally distributed across the classes. The final lists were revised to ensure that students who had been reported as experiencing social conflicts (e.g. bullying) in previous years were equally distributed amongst the different classes.

4.4.2. Experimental design

4.4.2.1. Randomisation

One of the characteristics of a true experiment is the random allocation of participants to the intervention and control groups. The correct assignment leads to the creation of two or more groups, which are, on average, equal with respect to the distribution (i.e. mean, SD) of observable and unobservable characteristics. Put differently, these random assignments allow the researcher to control for extraneous sources of variation in e.g. prior attainment or ability (C. Torgerson and D. Torgerson, 2003). Thus differences in outcome after the intervention can be attributed to the intervention, and not to pre-existing differences between the groups.

In education, randomisation at the individual level is often impossible for practical reasons. The school administration makes the randomisation of classes difficult because the classes cannot easily be divided during the academic year. Instead, randomisation might be carried out at the higher level units such as classes and schools (Moore, A. Graham and Diamond, 2003). However, C. Torgerson and D.

[2] Students in their last year of secondary education take a placement exam. The results of this exam determine the allocation of students to high schools and offers to enrol for the next level are conditional on the students' results. One of the assessed areas is reading comprehension.

[3] Although several attempts were made to gain access to the results of this test, no permission was granted. Information about the allocation procedure was provided by the psychologist responsible for the application of the test and the subsequent allocation of students.

Torgerson (2003) state that group randomisation will only show changes at the group level and not for every child. That is, the interactions between individuals in a group might have some effect on the intervention, or improvements might result from the relationship between individuals in the same group. In order to avoid these effects, Moore, A. Graham and Diamond (2003) propose a pragmatic approach to randomisation in educational studies. In this approach, the intervention is delivered under conditions as similar as possible to those found in real life. This means, for example, involving the existing teachers in the project. Moore, A. Graham and Diamond (2003) also advise supplementing the quantitative assessment of the effect of the intervention with qualitative research in order to understand better the underlying causal mechanism and which outcomes may be due to the remaining differences between the characteristics of the participants across the groups. Due to the school regulations and practicalities beyond the researcher's control (i.e. scheduled activities and post-intervention training delivered by teachers), it was impossible to randomise students at the individual level [4]. Instead, the five classes were randomly allocated to either the strategy or the control group (see Section 4.6.1).

4.4.2.2. Quasi-experiments

A true experiment is *"a study in which an intervention is deliberately introduced to observe its effect"* (Shadish, Cook and Campbell, 2002, p. 12). The aim of an experiment is to test a hypothesis that has been formulated based on observation.

An experimental design is a systematic approach that guides researchers' decision process when setting up an experiment. Depending on the characteristics of the design, the implementation of an experiment can be considered a true experiment or a quasi-experiment. The experiment is generally carried out with one or more control and experimental groups (here 'strategy') which should be *"carefully selected and sufficiently large for conclusions to be drawn about the larger population"* (Pring, 2004, p. 36). The present study, for example, aimed to introduce the direct teaching of reading comprehension strategies in order to observe the effect on the students' reading comprehension. An important component of a true experiment is the use of random processes to allocate the participants to the groups. The use of random procedures creates two or more similar groups with respect to observable and unobservable characteristics. The differences between the groups at baseline, that is

[4] Several attempts were made during the analysis of the results to minimise the potential bias due to a lack of randomisation at the level of the individual.

before the implementation of the intervention, may hamper the correct interpretation of the observed results. By ensuring the random allocation of students to the control and experimental groups, one can insure that the outcome differences at the end of the study are due to the intervention. True experiments are therefore the 'gold standard' in many areas of science, such as physics or medicine (e.g. C. Torgerson and D. Torgerson, 2003).

Quasi-experiments differ from true experiments due to the lack of random assignment. This allows differences between the control and the treatment group to persist after the allocation. Therefore, it cannot be ruled out that the differences in outcome at the end of the intervention are due to pre-existing group differences rather than the intervention itself. These confounding effects could provide alternative explanations for the observed effect and raise concerns over whether it is causal. Due to the challenges faced in educational settings to randomise students (McGowan, 2011; Moore, A. Graham and Diamond, 2003), quasi-experiments are commonly implemented in this area.

Quasi-experiments with the control group and pre-test are called the 'non-equivalent comparison groups design' (Shadish, Cook and Campbell, 2002). Having an untreated comparison group in a quasi-experiment is advantageous because it makes it easier to compare threats to validity, such as history, maturation, or retesting effect. Although the lack of randomisation can result in non-equivalence between groups, this kind of design allows the identification of observed differences between groups before the intervention, which can then be taken into consideration during the interpretation of the results. It is important to explore and be aware of any pre-existing differences between the groups before the intervention.

Shadish, Cook and Campbell (2002) state that this initial difference could be the result of 'selection bias' in some designs (with some element of voluntary participation) and suggest that this disadvantage of the quasi-experiment design could be avoided by adding design elements such as blinding the researchers and participants to the assignment, the use of multiple post-tests to determine the pattern of evidence about the effects, and the randomisation of higher-level units (e.g. classrooms or schools). To strengthen quasi-experiments, they suggest identifying threats of internal validity, collecting additional data on confounding variables (e.g. pre-test scores, information on students' socio-economic background), including additional control groups, and pattern matching.

A different source of bias is the 'intention to treat'. This means that researchers include in the analysis only participants who completed the intervention. However,

this approach has been shown to have the potential to bias the results because participants who decide to drop out of the study might have significant reasons for this. These reasons may have important consequences for interpreting the results and drawing implications and may, in themselves, be a source of bias (C. Torgerson and D. Torgerson, 2003). In order to address this issue, participants were only excluded from this study if the reason for their exclusion was judged to be unrelated to the intervention (e.g. medical reasons, moving to another city). Furthermore, the role of non-attendance was explicitly explored in the analysis of the influence of confounding variables.

4.5. Research instruments

Five instruments were used to collect data in order to address the five research questions (Figure 4.1): word-level study (WLS), reading comprehension tests, metacognitive assessment instrument (MAI), think-alouds, and background questionnaire. All tests were piloted with a different sample of students of the same age and educational level as the participants in the current study. The pilot took place in the same school approximately eight months before the start of the intervention. The details of this piloting are provided in Section 4.8. The following sections explain the design and validation of the instruments.

4.5.1. Word-level study

The first instrument was used to assess whether the students included in this study had problems with low-level processes. The word-level study comprised three tasks and aimed to assess the students' ability to decode, identify and comprehend Spanish words. Although this study focuses on higher-level processes, it is important first to ascertain the students' readiness with regard to the basic reading skills, i.e. that they are able to recognise graphemes, phonemes and the lexis of the Spanish language. Previous research findings suggest that poor comprehension might be caused by problems at both the lower and higher levels of the reading process (see Chapter 3). If decoding problems were found, this may suggest that trying to teach reading strategies in order to affect the higher-level processes may be inappropriate. In other words, problems with low-level processes, particularly if they differed between the intervention and the control group, would possibly be confounding the estimated effect of teaching reading strategies on comprehension. Identifying problems with

word level recognition may suggest that low levels of comprehension should first be addressed by improving basic word recognition, and perhaps subsequently through metacognitive strategy instruction, where necessary. In contrast, if no problems with low-level processes were detected, this would make it a) more appropriate to try to improve reading via metacognitive strategy instruction and b) more likely to identify the causal effect of strategy instruction on reading comprehension.

4.5.1.1. Word Identification

The word identification task identifies phoneme-grapheme correspondence problems. The students were presented with a list of Spanish words, which they read out aloud. The Spanish language has an extensive lexis of approximately 88,000 words (Real Academia Española, 2001). To ensure that the word identification task covered a representative part of the Spanish language, a list was created, taking into consideration the number of word syllables and word frequency. The final list contained 35 Spanish words and is presented in Appendix C.1. Words were stratified and randomly sampled from the 'Corpus del Español' (Corpus of Spanish) (Davies, 2002). The words were first grouped according to the number of syllables and then according to their frequency. The final list contained 14 words (40%) with two to three syllables, 14 words (40%) with four to five syllables, and 7 words (20%) with six or more syllables. The frequency of the words was also considered in order to test whether the students had a better knowledge of high frequency words than low frequency words. The list includes 8 high, 13 medium, and 14 low frequency words, distributed evenly between the different syllable lengths. The list was piloted on 17 students (see results in Section 4.8.1).

4.5.1.2. Word comprehension

The word comprehension task assessed whether the students' vocabulary comprehension is comparable to the norms for their age. This allowed the researcher to verify whether vocabulary problems might be the cause of comprehension problems. Word comprehension was assessed using the vocabulary and similarities subtests from the Spanish revised version of the Wechsler Intelligence Scale for Children (WISC-R)(Wechsler, 1974; Rodriguez and Prewitt Diaz, 1990). The WISC-R test is an intelligence test that can be administered to children aged between 6 and 16 years old. The test generates a composite score for overall cognitive ability, which can be divided into four individual sub-scores: i) the verbal comprehension index (VCI), ii)

the perceptual reasoning index (PRI), iii) the processing speed index (PSI), and iv) the working memory index (WMI). The verbal comprehension index (VCI) measures verbal concept formation, i.e. the child's ability to reason verbally. It includes five subtests: vocabulary, similarities, comprehension, information, and word reasoning. For the current study, the subtests of vocabulary and similarities were chosen to assess the students' knowledge of Spanish vocabulary [5]. The Spanish version of the WISC-R test has been validated and the discriminatory power and test-retest reliability have been demonstrated in a Spanish speaking population across various age groups (Rodriguez and Prewitt Diaz, 1990; Gass, Demsky and Cameron Martin, 1998).

In the vocabulary subtest, the students were asked orally to define 32 words (see Figure 4.2 and Appendix C.2). Each correct definition was awarded 2 points, so that a total of 64 points could be achieved. Partial scores (i.e. 1 point) were awarded for incomplete definitions. In the similarities test, the students were presented with 17 pairs of words, then asked to explain how the words were alike (see Figure 4.3 and Appendix C.3). Answers were awarded with one or two points, with a possible total score of 30 points. The final scores were compared to standardised values for the corresponding age groups provided in the appendices of the WISC-R application manual (Rodriguez and Prewitt Diaz, 1990). A copy of the actual tests can be consulted in Appendix C.4.

knife	dangerous

Figure 4.2.: Example of the WISC-R vocabulary subtest

How are the words **wheel** and **ball** alike?
How are the words **elbow** and **knee** alike?

Figure 4.3.: Example of WISC-R similarity subtest

4.5.2. Reading comprehension test

The comprehension test sought to measure and compare the students' reading comprehension before and after the intervention in order to test whether the strategy

[5] The comprehension subtest was not included as an outcome measure because it consisted of questions. The students had to explain how they would act if they found themselves in a similar situation (e.g. What would you do if you found a wallet belonging to someone else in a shop?).

instruction affected their reading comprehension. The students were assessed on three occasions: at the beginning of the study (0 months), immediately after the intervention (2 months) and four months after the intervention ended.

The test was designed using materials taken from previous PISA assessments (2000, 2003, 2006, 2009) and from three study guides designed for Mexican students: the Bachelors' Guide (Guía para Bachillerato) (*Guía para Bachillerato* n.d.), the Magic Guide (Guía Mágica) (Robles Robles and Olvera, 2002), and Guide 11 (Guía 11) (Carrillo López, Gil Chaveznava and Morales López, 2009) [6]. In order to strive for ecological validity, the tests were designed to elicit and measure particular metacognitive reading strategies. The questions were 'matched' with the reading strategies that were taught.

An inherent problem with the repeated assessment of individual students is improvement due to adaptation to the test or memorization of specific items; known as the 'test effect' (L. Cohen, Manion and Morrison, 2011). To reduce the test effect, three versions of the test were designed (A, B, and C). Each version consists of five texts, each containing three or four question items (i.e. multiple choice and open questions). The questions assess the three aspects evaluated in PISA: access and retrieve, integrate and interpret, and reflect and evaluate. To ensure validity and equivalence among the three versions, the tests were piloted and the results analysed by carrying out an item analysis and analysis of variance (ANOVA) (see Section 4.8.4).

The nature of the strategies tested was matched across the different versions, to ensure each test required the use of the same strategies. Furthermore, the length and readability of the texts as well as similarities between the topics were compared across the three versions (Tables 4.1 and 4.2). To do so, readability indices were calculated using the Huerta Reading Ease formula (Fernández Huerta, 1959). This formula is appropriate for calculating the readability indices of texts written in Spanish because it integrates the average number of syllables and sentences [7]. The results indicated that the readability across texts was adequate for the students' age and approximately equal across texts (+/- 5 points on a scale of 0 - 100).

[6] The study guides contain activities related to the different school subjects and are tailored to students in their last year of secondary education.

[7] For the English texts, the readability indices can be calculated using the Flesch Reading Ease formula (Flesch, 1948). However, the words and phrases in Spanish are longer than the English ones. Therefore the Flesch formula does not provide accurate results.

	Word Count			Readibility		
	Version A	Version B	Version C	Version A	Version B	Version C
Text 1	544	600	605	69.6	64.9	73.7
Text 2*						
Text 3	300	334	314	75.8	71.5	67.7
Text 4	382	393	390	79.6	73.5	78.6
Text 5	264	274	328	73.5	79.8	75.6

(*) The second text in the test is a chart and it was therefore not possible to calculate the word count or readibility index.

Table 4.1.: Readibility index and word count of the reading comprehension tests

Text	Description (in Spanish)
Version A	
Text 1	La tecnología crea la necesidad de nuevas leyes
Text 2	México hacia el 2000
Text 3	Mueren tres mil peces por falta de oxígeno en un lago de Aguascalientes
Text 4	Las abejas
Text 5	Mi primer amor
Version B	
Text 1	Bioética
Text 2	PLAN Internacional
Text 3	Siéntase bien en sus zapatos deportivos
Text 4	Delfines
Text 5	La casa encantada
Version C	
Text 1	Las armas científicas de la policía
Text 2	Iberoamérica Crece
Text 3	El te podría mejorar la memoria
Text 4	Las nubes
Text 5	La sangre de Medusa

Table 4.2.: Topics of texts in the reading comprehension tests

As noted above, the students' reading comprehension was assessed on three

occasions. One version of the reading comprehension test (A) was used as the pre-test and the remaining versions served as an immediate post-test (B) and delayed post-test (C). The researcher administered the immediate post-test at the end of the first stage of the intervention and the delayed post-test four months later. Copies of the tests are presented in Appendix C.5 (version A), Appendix C.6 (version B), and Appendix C.7 (version C).

4.5.3. Metacognitive Assessment Instrument

Metacognition is defined as knowledge about and the regulation of cognition. Its influence on reading comprehension is generally accepted by the scientific community (e.g. Baker and Carter, 2008). Studies describing skilled readers have demonstrated that the use of metacognitive processes directs the readers' efforts towards successful comprehension (Garner and Kraus, 1981). Three distinct elements have been emphasised as being central to metacognition: planning, supervision, and evaluation (Garner and Kraus, 1981; Garner, 1987). One means of measuring readers' awareness and control of metacognitive strategies is through self-report instruments e.g. questionnaires and verbal reports. This kind of instrument allows students to express their perceptions about a topic freely, without feeling judged by the researcher. However, two of the disadvantages of self-reporting are that a) the answers reflect a subjective assessment of the awareness and control of metacognitive strategies, which may not correlate well with their actual levels of control, and b) the students may purposefully or unintentionally misreport (Young, 2005; L. Cohen, Manion and Morrison, 2011).

Researchers have noted the importance of developing a more objective scale that allows the measurement of the students' awareness and perceived use of strategies. Several instruments have been developed for this purpose, such as the Index of Reading Awareness (Jacobs and Paris, 1987), Reading Strategy Use (Pereira-Laird and Deane, 1997), Metacognitive Reading Awareness (McLain, Gridley and McIntosh, 1991), and the MSI Metacomprehension Strategy Index (Schmitt, 1990; Miholic, 1994). However, these instruments have been criticised for featuring only a small number of items and showing poor psychometric properties (i.e. dubious validity and poor test-retest reliability). Moreover, most of them focus on primary school students, which limits their usefulness for the assessment of middle or upper level students.

The metacognitive assessment instrument (MAI) used in the current study was designed to measure the students' perceptions of their strategy use as well as their

reported frequency of use. This self-reported instrument assessed strategy use, facilitating a comparison of the students' responses before and after the intervention. Two self-reported instruments served as the basis for the design of the metacognitive questionnaire used in the current study: the Metacognitive Awareness of Reading Strategies Inventory (MARSI), developed by Mokhtari and Reichard (2002) and the Escala de Conciencia Lectora, Reading Awareness Index (ESCOLA), developed by Jiménez Rodríguez (2004). Both instruments measure metacognitive functions; however, they were validated among different populations. MARSI was validated with English-speaking adolescents and adults whereas ESCOLA was validated with Spanish-speaking students (8 - 13 years old).

Designed to be used by adolescents and adult readers, MARSI assesses the *"students' awareness and perceived use of reading strategies while reading academic or school-related materials"* (Mokhtari and Reichard, 2002, p. 249). It is composed of 30 items that assess strategy use. These strategies can be grouped into three categories: global reading strategies, problem-solving strategies and support reading strategies. Each statement is measured using a five-point rating scale. The steps followed for its development and validation ensure high levels of validity and reliability (Mokhtari and Sheorey, 2008). This instrument was created based on an extensive review of the research literature on reading comprehension and metacognition, a review and analysis of the previous reading strategy instruments, discussions with experts in the area, and a factor analysis of the items. It was revised on several occasions and piloted with an extensive sample (N=825) to ensure high reliability (Cronbach α = 0.89) (Mokhtari and Reichard, 2002).

ESCOLA was created to assess reading skills and predict reading development in Spanish speakers. It is composed of 56 multiple-choice questions that are answered while the students perform a reading comprehension task. It was developed and piloted in Spain and Argentina. ESCOLA was designed to measure students' reading awareness and to assist with the development of reading programmes. Such programmes are tailored to solve reading problems detected from information gathered via ESCOLA (Fernández Lozano, Jiménez Rodríguez and Alvarado Izquierdo, 2010). ESCOLA measures three elements of metacognition (i.e. planning, supervision and evaluation) and their interaction with three variables: the person (the reader's self-perception and skills); the task (its difficulty and ambiguity); and the text (its characteristics). Planning is measured by items that assess readers' attitudes towards the test and the selection of reported strategies articulated at the planning stage to achieve the objective of the task. Supervising is measured by the level of readers'

attention while reading, the use of strategies to choose relevant information in the text, self-efficacy [8], and perseverance. Evaluation is measured by items that assess the control over reader's performance, access to the correct strategies, and recognition of comprehension success. ESCOLA showed good reliability in the Spanish and Argentinian samples (Cronbach α = 0.81 and 0.86). The results from the reliability tests showed that these results could be generalised to other samples. Convergence validity was tested by comparing ESCOLA with two comprehension tasks, a questionnaire answered by the students' tutors and a Spanish adaptation of MARSI developed by the authors of ESCOLA (Jiménez Rodríguez et al., 2009). The results from the comparison with the comprehension tasks indicated a low correlation with the first test (0.14 - answer questions) and the second test (0.25 - cloze test). The results from the comparison with MARSI showed that 51% of the variance of MARSI is explained by ESCOLA and that the two tests have a high correlation (0.72).

For the current study, statements from the two instruments, MARSI (in English) and ESCOLA (in Spanish), were translated and/or adapted to match the strategies taught during the intervention. Note that only statements related to the strategies taught in the current intervention were chosen and adapted and other statements were discarded. A list of 58 items was produced and two independent reviewers checked them for inconsistencies, readability and whether the items assessed the same constructs. Based on the reviewers' comments, the list was reduced to 50 items. Some questions were altered (negatively worded) to avoid mechanisation (L. Cohen, Manion and Morrison, 2011). This procedure helps to detect students who might answer the items automatically without thinking carefully about the questions asked. After the piloting (see Section 4.8.2), a final version of the MAI instrument was produced that features 48 items, each of which uses a four-point rating scale. By employing a four-point rating scale, the researcher tried to control 'the central tendency' phenomenon (L. Cohen, Manion and Morrison, 2011). It has been observed that the participants tend to choose the mid-point of five-point scales. The final version of the MAI is reproduced in Appendix C.8.

4.5.4. Verbal reports

Verbal reports (i.e. think-aloud protocols) were designed to assess whether the students really use the strategies they reported in the MAI, i.e. to corroborate the

[8] Self-efficacy is defined as knowing and controlling the tools that help the student to comprehend a text.

results yielded by the self-report questionnaire. In any self-assessment questionnaire, there is a risk that the respondents will intentionally or unintentionally misreport the use of reading strategies (L. Cohen, Manion and Morrison, 2011). In this study, three main problems were foreseen: a) overrating; i.e. students reporting an overly high use of a specific strategy, b) students reporting falsely that they know and use the strategies when this is not reflected in their actual use, and c) students confusing strategies and therefore reporting the use of the wrong strategies. Carrell (1989), for instance, found that, despite the fact that students possess strategies, they sometimes report using different strategies from those that they actually use. Therefore, it is important to corroborate the self-reported use of strategies during comprehension tasks to assure the reliability of the results.

Verbal reports are used to elicit the students' thoughts while engaged in a given activity (Young, 2005). The students are asked to think aloud about the processes occurring in their mind during the activity. This allows the researcher to record the actual steps that the students follow while performing a task. Data can be collected while the students perform the task (i.e. concurrent protocols) or after it has been performed (i.e. retrospective protocols) (Bowles and Leow, 2005). The main aim of think-aloud protocols is to support and verify data that have been collected by other means (e.g. observations, interviews). The participants are provided with engaging activities which prevent them from providing what they think is a desirable response. Also, in the case of concurrent think-alouds, the immediate report of thoughts reduces memorization problems with the later recall of information (Young, 2005). However, this type of data collection requires students to externalise their thoughts by talking about cognitive processes; therefore, potential disadvantages such as reactivity, the participants' poor verbal ability, and data validity have been noted (Young, 2005). Reactivity refers to the participants' reactions while performing two tasks (i.e. thinking and reporting); it supposes that reporting might alter the participants' cognitive processes. Also, some students may not possess the appropriate verbal skills to describe these processes. Therefore, long periods of silence might occur and impede the efficient collection of data. Finally, automated processes might not reach the conscious level and therefore go unreported.

The verbal reports were collected before and immediately after the intervention to allow a comparison of students' performance. The students were told to read a text and perform the comprehension task (answer questions). They were also told to speak aloud and state whatever they were thinking about, including memories that the text might have brought to their mind. They were asked to notify the researcher if

and why they had trouble understanding the text or whether they found any mistakes in it. While the students were reading, they were prompted by the researcher to share their thoughts each time they either fell silent for more than 30 seconds, when they seemed troubled, or when they read backwards. The students' responses were recorded using a tape recorder for later decoding and analysis.

To ensure the appropriate use of verbal reports, the students received instructions and a brief practice in the process of think-alouds before being presented with the actual text. Three analogy tasks were extracted from the study guides. The exercises provided a statement containing two related words. The students were asked to choose one of the five pairs that contained words with a similar relationship as the pair given in the statement (see example in Figure 4.4). The researcher used one of the analogies to show the students how verbally to describe their thinking while solving the task. Then, the students were given the opportunity to answer two more analogies. Feedback was provided, and when the students forgot to talk, they were prompted. After three examples, they were given the actual task.

A triangle is to an architect as a ...

1. patient is to a doctor
2. book is to a writer
3. compass is to a material
4. *scalpel is to a surgeon* (correct answer)
5. animal is to a veterinarian

Figure 4.4.: Example of analogy task in verbal report

As part of the verbal reports, students were asked to read texts and report their thinking. The texts were taken from a weekly Mexican magazine, 'Dia Siete'. The magazine is sold as part of a national newspaper and includes articles written by nationally and internationally recognised authors. The variety of topics (e.g. politics, science and economics) and formats (e.g. tables, narrative and interviews) provide a rich source of texts on current topics. Many of these topics were unfamiliar to the students. Three texts were used for the verbal reports. These are reproduced below and in the Appendices C.9 and C.10.

4.5.4.1. Pre-test version

The pre-test version of the instrument was relatively short (436 words). The topic of the text was nakedness and the way it is perceived. The author described nakedness as a natural state and provided examples and evolutionary reasons for this. He criticised the embarrassment that human beings feel when naked. He introduced the text by comparing the nakedness of the statues in parks with human nakedness. The text was chosen because several topics are described within a short text which requires the use of inference, background knowledge and context to understand the meaning of unfamiliar words and so fully comprehend the text. The task targeted the use of the following strategies: identify and retrieve direct information, understand the main idea of a paragraph, paraphrase it, and make inferences to solve the tasks. The actual text and an approximate English translation of the questions are provided in Figure 4.5 and Table 4.3.

La desnudez del mármol

Siempre me han dado frío las estatuas. Sobre todo las de mármol. De cuando en cuando uno las halla en mitad del parque y un escalofrió se apodera de nuestra conciencia. “Pobrecita petacona a la intemperie, yo con mi bufanda y ella buscando no se que bajo la luna de invierno.” La desnudez fue un acertado invento de la especie, de cuando no necesitábamos todo este bagaje de última moda, a lo Armami, y peludos y feos deambulábamos tan quitados de la pena igual que los orangutanes de Malasia. Pero habría de llegar la expulsión bíblica del Paraíso para darnos cuenta, como si salidos del sauna, de nuestra desnudez repentina luego que inventamos ese concepto terrible llamado “vergüenza”.

Los antropólogos no se ponen de acuerdo, pero algo bueno hubo cuando perdimos nuestra pelambre. Desmond Morris, en su legendario ensayo El mono desnudo, argumenta que nos quedamos en cueros debido a que de ese modo pudimos librarnos de las pulgas que infestaban nuestra piel cuando habitábamos en las cavernas. José H. Reichholf, en su deslumbrante libro La aparición del hombre (traducido a 20 idiomas), establece una tesis más original. Afirma que vamos desnudos por los caminos de Dios debido a que sólo de este modo podemos sudar lo suficiente para refrigerar el cuerpo, pues nuestra especie, la de los Homo sapiens, se distinguió de los otros simios porque fuimos (y somos, ¿verdad Ana Guevara?) bípedos corredores. Según el estudioso, logramos desarrollar nuestra capacidad cerebral gracias al fósforo que obteníamos de la carne que devoramos, durante milenios, compitiendo con los buitres. Esto es, que fuimos una especie carroñera corriendo tras la primera señal de un animal muerto en las estepas de lo que hoy es Tanzania y sudando, solamente sudando, es que nuestro cuerpo aguanta correr durante una hora.

El hecho es que vamos lampiños por la vida sin saber a bien qué hacer con nuestra desnudez. Nacemos desnudos y nos desnudamos para amar, desnudos nos bañamos y desnudos (o casi) nos adentramos en el océano para practicar la natación, habida cuenta de que un cuerpo desnudo equivale a un grito primigenio: ya sea en la playa de Biarritz o en el vestíbulo del Palacio Legislativo, un hombre o una muer desnuda es una bofetada a nuestra conciencia (tan pudorosa), ahita de calzones, pijamas y gabardinas.

En Sinaloa les llaman “bichis” y en Argentina “andar en pelotas.” Desnudos, encuerados, desvestidos. La piel que cada ano exige su dosis de vitamina E, una semana bajo el sol de Acapulco, y una noche de caricias para no convertirnos en frías estatuas de mármol refundidas en lo más oscuro del parque.

Figure 4.5.: Text used in verbal protocol (pre-test)

Five incongruences were included in the text: a) the word order was changed in two sentences, b) two words were modified in order to create semantic confusion while reading the sentence, and c) a word was misspelled (Table 4.4). The aim of introducing incongruences was to assess whether the students were reading the ‘form’ of the text carefully (i.e. applying some degree of effort and attention to decoding), whether they were able to identify these incongruences and whether they were able to overcome the interference caused by these incongruences. The appropriate identification would suggest that they were reading with a sufficient high level of effort and attention (Markman, 1979; Zinar, 2000).

1.	*Menciona las dos posibles causas de la desnudez humana.* (Give two reasons why human beings are naked.)
2.	*¿Qué exigencias tiene nuestra piel?* (What does the skin need?)
3.	*Señala el párrafo que expresa que la desnudez es el estado natural del hombre.* (Highlight the paragraph that says that nakedness is the natural state for human beings.)
4.	*Además del desplazamiento en dos pies, ¿Qué otro factor nos alejó de los demás simios?* (In addition to walking upright, give another characteristic that differentiates human beings from other primates)

Table 4.3.: Test questions in verbal protocol (pre-test)

4.5.4.2. Post-test versions

A modified version of the pre-test text was used for the verbal protocols after the intervention. The original text was given to a Mexican literature teacher who was told to amend it in order to generate an alternative version. She used synonyms and paraphrased some sentences in the text but kept the main ideas and data (Appendix C.9). The questions about this text are presented in Table 4.5. In addition, a second text was used to elicit strategies different from those needed to solve the tasks for the first text (Appendix C.10). This second text relates a legend concerning how bats used to have colourful feathers and lost them. The tasks for this text required the students to perform a more in-depth analysis by evaluating and sharing their thinking about the text (Table 4.7). Incongruences in the sentences and words were also added to both post-test texts to verify the students' monitoring (Tables 4.6 and 4.8).

Incongruence	Correct version	Amended version	English
Change of word order	Los antropólogos no se ponen de acuerdo [...]	Los antropólogos de acuerdo se ponen no [...]	Anthropologists do not agree [...]
	Nos adentramos en el océano para practicar la natación	Nos adentramos en la natación para practicar el océano	We go into the ocean to practice swimming
Semantic confusion	cuerpo	cerco	Gate instead of body
	piel	miel	Honey instead of skin
Misspelled word	mujer	muer	woman

Table 4.4.: Incongruences in verbal protocol (pre-test)

1	*Señala la oración donde se asegura que la desnudez es el estado natural del hombre.* (Find the sentence that states that nakedness is the natural state for human beings)
2	*¿Qué necesita nuestra piel para no ser como la de las estatuas?* (What does the skin need to be like in order to differ from that of statues?)
3	*Según el texto, ¿Qué tuvo que suceder para que sintiéramos vergüenza de nuestra desnudez?* (According to the text, what had to happen that made humans feel embarrassed by their nakedness?)
4	*En esta lectura se mencionan dos características que nos hacen diferentes de los primates, ¿cuáles son?* (Find the two characteristics that make primates and human beings different)

Table 4.5.: Test questions in verbal protocol (post-test version 1)

Incongruence	Correct version	Amended version	English
Change of word order	[...] al quedarnos sin pelambre dejamos en ella los bichos que nos invadian viviendo en las cavernas.	[...] al quedarnos pelambre sin dejamos en ella los bichos que nos invadian viviendo en las cavernas.	When we lost the [fur] we lost the flies we had while living in caves.
	[...] como le dicen en Argentina	[...] como dicen le en Argentina	As it is called in Argentina
Semantic confusion	años	arcos	Archs instead of years
	hora	obra	Play instead of hour
Misspelled word	invento	ivento	invention

Table 4.6.: Incongruences in verbal protocol (post-test version 1)

1	*¿Qué titulo deberia tener por su contenido?* (Give the text a title based on the story)
2	*¿Cuál de estas palabras califica mejor la actuacion del murcielago? Tonto, presumido, malvado, envidioso. ¿Por qué?* (Which of these words best describe the bats performance? Fool, conceited, evil, envious. Why?)
3	*¿Qué opinas del comportamiento de las demas aves? Fundamenta tu opinion.* (What do you think about the other birds behaviour? Justify your answer.)
4	*¿En que oracion se expresa que Dios manda llamar al murcielago?* (Which sentence states that God called the bat)
5	*Menciona una frase corta que pueda ser la moraleja de este relato.* (Give a short phrase that could sum up this story.)

Table 4.7.: Test questions in verbal protocol (post-test version 2)

Incongruence	Correct version	Amended version	English
Change of word order	Ya no tenía ninguna pluma	Tenía no ya ninguna pluma	He did not have any feather
	Una parvada de pájaros	Una pájaros de parvada	A flight of birds
Semantic confusion	aves	Naves	Ship instead of bird
	plumas	uñas	Nails instead of feathers
Misspelled word	tarde	tade	afternoon

Table 4.8.: Incongruences in verbal protocol (post-test version 2)

In order to verify the validity of the texts, they were given to two external reviewers: a Mexican university student and a Mexican literature teacher. They were told to read the text, perform the tasks, and report any incongruence they could find. Both reviewers identified four of the five incongruences in each of the three texts. After prompting them for a further mistake, they were able to find the misspelled word. Therefore, in the final version of the text used in the main study, the misspelled words were changed to more frequently used words to make them more likely to be noticed.

4.5.5. Background questionnaire

The background questionnaire was given to all students at the beginning of the intervention. It consisted of three parts. First, the students were asked about their general characteristics (e.g. age, gender) as well as for information regarding their socioeconomic circumstances (e.g. parents' occupation, number of rooms in their house). The questions in the second section focussed on the students' reading habits and attitudes towards reading (e.g. time spent reading at home, type of books, etc.). Finally, the third section elicited information regarding reading activities in the classroom. Information from this questionnaire was used to test whether the students' characteristics were equally distributed across both groups. This information was also used to investigate any relationship that these factors might have with the comprehension scores by employing them as covariates in an analysis of covariance (ANCOVA). A copy of the background questionnaire can be found in Appendix C.11.

4.6. Teaching procedure

4.6.1. Sampling method

Five intact groups were randomly allocated to the strategy and control group. Randomisation was achieved by tossing a coin. Three classes with a total of 103 students formed the strategy group. The students in this group received direct instruction on the use of reading strategies. Two classes with a total of 74 students formed the control group. The students in this group received normal instruction.

Before the intervention began, all participants took one version of the reading comprehension test to assess their comprehension before the intervention. To counterbalance the three versions of the test, both groups were now randomly divided into three subgroups. Each subgroup took a different version of the three reading comprehension tests as a pre-test.

A subsample of 62 participants was randomly chosen from each class to participate in the word-level study and the think-alouds. This was done by ordering the students within each class alphabetically and choosing every third student for inclusion, i.e. student one, four, seven, and so forth. This procedure yielded 19 students from the control group and 43 students from the strategy group. The researcher met each student individually in the library of the school where they took the tests.

4.6.2. Criteria for choosing the strategies

As the review of the literature in Chapter 3 illustrates, while there is substantial evidence for the multiple benefits of reading strategies, there is little agreement about which strategies are the most effective or about the number of strategies that should be taught. Block and Duffy (2008), for example, list 45 strategies studied over a period of 10 years. In order to enhance the students' reading comprehension, researchers have repeatedly developed instructional approaches (see Chapter 3) that aim to teach different strategies. As a result, studies are conducted under different conditions (i.e. place, participants and research methods) and classified differently in different studies. Furthermore, the same strategies are named differently across the studies, making it difficult to summarise their activity through meta-analysis.

To date, research has shown that the teaching of multiple strategies yields better results than the teaching of single strategies in isolation (see Chapter 3). However, the lack of agreement in this field of research often represents a problem when designing a new study. In order to make an informed decision about the strategies to be taught in this study, a pre-specified set of criteria were followed. First, strategies identified as effective for improving students' comprehension in studies between 1970 and 1990 were identified. Only strategies that had been further studied, in at least two studies up to 2000, were retained. Next, a search was carried out to identify experimental studies published between 2000 and 2009 that had directly taught the remaining strategies in real-world settings (i.e. schools). The results served to redefine the list (Appendix B.1) which was then reviewed to ensure that it contained the nine essential strategies suggested by Block and Duffy (2008) and the strategies used by good readers (Houtveen and Grift, 2007; Pressley, 2000). Sixteen strategies were chosen as being the most significant and useful for adolescent students working with expository texts. They are listed in Table 4.9.

4.6.3. Teaching intervention

The project lasted 6 months in total and was divided into three stages. Each stage was approximately two months long.

4.6.3.1. First stage: Initial testing and researcher interventions

During the first stage of the intervention, both groups worked on the same texts and carried out the same tasks (e.g. answer questions). The researcher delivered

Strategy	Description	Introduced in week ...
S1	Pre-reading strategies & goal setting	1
S2	Use of prior knowledge	1
S3	Self-questioning	6 and 8
S4	Making inferences	1 and 9
S5	Visualising	9
S6	Rereading	2
S7	Reading forwards/backwards	2
S8	Change of reading speed	6
S9	Use of context	1
S10	Use of external sources	11
S11	Summarising	10
S12	Main ideas	3
S13	Paraphrasing	3
S14	Monitoring and supervising	4 and 6
S15	Evaluating	6
S16	Strategy awareness	5

Table 4.9.: List of strategies taught as part of the intervention

teaching to both groups in order to avoid teacher bias. The students in the strategy group were directly instructed in the use of reading strategies. The students in the control group, also with the researcher, read the same materials as the strategy group but using methods that are more typically used by teachers in secondary education (Treviño et al., 2006).

During the first week of the academic year, the students took the reading comprehension pre-test and completed the background questionnaire and the MAI pre-test. A subsample of students was tested with regard to their basic reading skills and their actual use of strategies using the verbal reports. The students were told that their answers would help to tailor the workshops to their needs; hence, the importance of providing honest answers. After the initial assessments, all of the participating students met the researcher for a 50-minute session twice per week for two months. In total, the researcher taught the students for 12 sessions (600 minutes).

Due to practical reasons and the internal organisation of the school, the working groups for both conditions consisted of 15-17 students. As mentioned before, the

intervention was included in the school activities as part of the afternoon workshops which lasted two hours. However, it was considered that a two-hour lesson devoted to reading activities would exhaust the students, especially after all the other academic activities of the day. Hence, each group in both conditions was divided into two smaller 'working subgroups'. The students of each subgroup were randomly allocated to either the reading workshop with the researcher or to working as normal with the Spanish teacher. After the first hour, the subgroups swapped classrooms. Working with smaller groups ensured a good control of the group and allowed closer, more personalised monitoring of the students during the first few weeks.

4.6.3.1.1. Strategy instruction

The strategy instruction aimed to promote the three types of knowledge described in the learning theory literature: declarative, procedural and cognitive (J. R. Anderson, 2009; Paris, Lipson and Wixson, 1983) (see Chapter 3). To make the concept of the strategies accessible and attractive to students, the strategies were introduced using a metaphor - each strategy represented a working tool. The students in the strategy group were told that they would be working with a toolbox, which contained tools (or strategies) that would help them solve their reading problems. There was also a prompt in the form of a glove, which symbolised five cognitive strategies (i.e. goal setting, use of context, prior knowledge, visualise, and making inferences) and three metacognitive strategies (i.e. monitor, evaluate, supervise) (Figure 4.6). This prompt served to remind students of how to use the strategies by providing them with steps to follow.

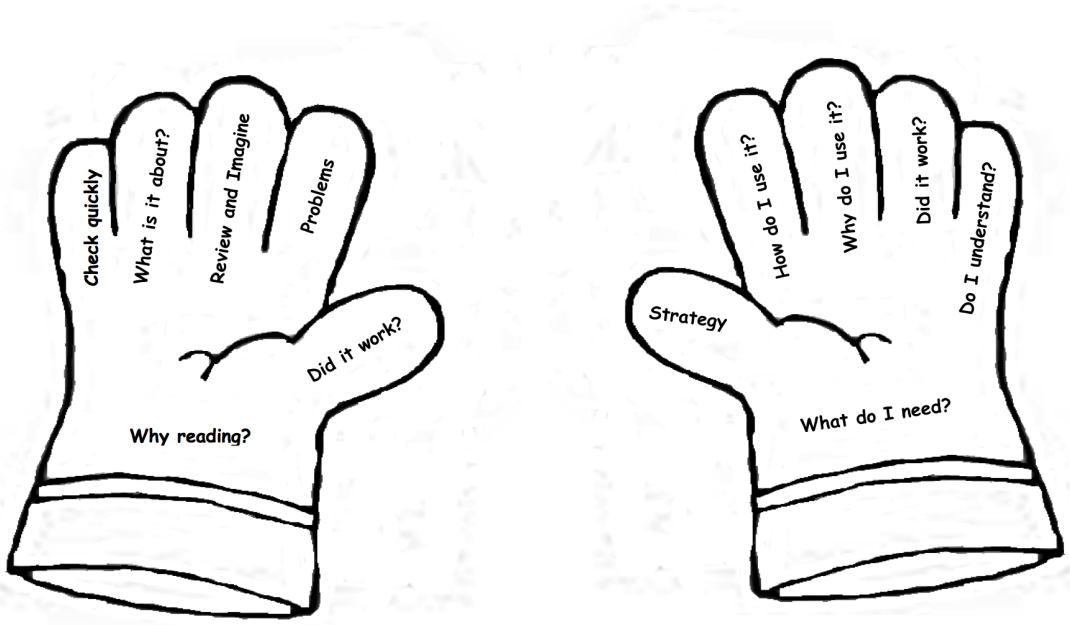


Figure 4.6.: Prompts used during the strategy instruction

The front of the hand included six steps:

1. *Why to read?* (¿Por qué leer?) - The students were told about the importance of setting goals for their reading and thinking about the purpose of their reading before starting.

2. *Quick review* (Reviso rápido) - The students were taught to perform a quick scan of the text and use clues such as titles, subtitles, drawings, charts, and any highlighted information to create an idea of what the text is about.

3. *What is it about?* (¿De qué se trata?) - Using the information from the previous step, the students drew inferences about the meaning of the text.

4. *Review and imagine* - The students were instructed to read half of the text and then stop to think what had happened in the text so far. They were told to try and create images in their head while reading. Also, they had to try to guess what the second part of the text was about. This step included three strategies: visualise, make inferences, and monitor.

5. *Problems* (Problemas) - After reading the text for the first time, the students were trained to stop and identify comprehension problems. They were told to look for parts, phrases or words that had caused them problems. In order to solve these problems, they needed the 'tools from the box'. Whenever a student was faced with a comprehension problem, they would turn the glove around, choose a strategy and follow the steps to apply the strategy successfully.

6. *Was it useful?* (¿Me sirvió?) - After choosing and applying one of the strategies in the box, the students were told to evaluate whether the strategy had been useful and also whether the text had been useful in achieving the initial goals.

The steps on the reverse helped to guide the students' use of metacognitive strategies and the three types of knowledge (see Chapter 3). This side also consisted of six steps:

1. *What do I need?* (¿Qué necesito?) - The students were trained to think about the difficulty they were facing.

2. *Strategy* (Estrategia) - The students chose one of the strategies contained in the tool box (declarative knowledge).

3. *How to use it?* (¿Cómo se usa?) - The students had to think about the process or how that strategy works (procedural knowledge).

4. *Why do I use it?* (¿Por qué la uso?) - The students were taught to become aware of the reasons for using a strategy and when to use it (cognitive knowledge).

5. *Did it work?* (¿Sirvio?) - The students needed to evaluate whether the chosen strategy had helped them to solve the problem, i.e. they were supervising their comprehension.

6. *Do I understand the text?* (¿Entiendo el texto?) - The students were taught to evaluate their comprehension after applying the strategy. They were also creating questions themselves about the text.

The remaining strategies were introduced as problem-solving tools which were inside the box. The students had six strategies to choose from: use of prior knowledge, visualisation of information, rereading, use of context, use of external information sources, and summarise. It was explained that, with time and practice, they would learn which were the most useful strategies for them. Every time a strategy was

introduced, the students learned the name of the tool (declarative knowledge), the steps to follow in order to use it (procedural knowledge), and when and why the tool was needed (cognitive knowledge) (J. R. Anderson, 2009; Paris, Lipson and Wixson, 1983). Through practice students received opportunities to apply the strategies in real contexts (automatisation).

The instruction was designed in line with the characteristics of previous studies that proved successful (see Chapter 3). Four characteristics were common to the approaches: a) teacher explanations, b) practice in real contexts, c) monitoring and additional modeling, and d) feedback to the learners. A strategy or group of strategies was introduced and modelled at the beginning of every week (Table 4.9). The introduction of the strategy included the elicitation of previous knowledge to find out whether the students were already familiar with the strategy and whether they knew how and when to use it. After the strategy had been modeled, a brief explanation was given to help the students to realise the importance of it and the situations in which they could use it (declarative knowledge). During the remaining days, the groups were organised into smaller groups of 3 to 4 students (but still within the group of 15-17 students) to realise the benefits of cooperative learning. The students were involved in situations where they would be able to practise the strategy using task-based activities (proceduralisation), thereby leading to automatisation. The students were instructed to discuss and share ideas (i.e. modeling strategies) in order to solve their comprehension problems. One member of each group served as a monitor to guide the group's work. The students in charge of each group ('supervisors') assessed the work of their peers by eliciting the strategy or strategies that could be used in each situation and ensuring that their peers followed the steps taught as part of the proceduralisation. The practice developed from very easy and controlled activities to more complex and independent activities. A sample of the tasks can be found in Appendix C.12.

4.6.3.1.2. Control group activities

The students in the control group read the same texts and worked with the same tasks as the strategy group. These students focused on discussions about the topics covered in the texts and followed the procedures normally used with their teachers: read the text, underline and discuss unknown or difficult words, listen to the teacher's explanation of the topic. They were told how the workshop would be structured and about the benefits of reading different kinds of texts as well as the importance of the

follow-up activities. It was explained that discussions would take place after reading each text in order to help them to understand the texts.

The activities designed for the control group were based on observations made by the researcher during the pilot study, informal talks with the teachers and a study conducted in Mexico which reported the instructional approaches used by teachers in primary education to foster the students' reading comprehension (Treviño et al., 2006; INEE, 2006a). In this study, first, fourth and fifth grade teachers answered questionnaires to measure the frequency of the utilisation of various reading activities. The responses showed that 78 per cent of the teachers work in schools that develop activities to promote reading among students. The teachers reported promoting activities such as answering questions related to texts, the choral reading of texts, and copying phrases or words from the text. Moreover, the results showed that the teachers emphasised the importance of developing word recognition and low-level reading skills but not higher-level skills. The responses also indicated that the teachers assign silent reading activities, especially when absent from the classroom. It was also found that the time allocated to reading activities is limited (i.e. less than an hour per day). Eighty per cent of the teachers reported participating in one-day training sessions on comprehension strategy development. However, the study does not report any instruction in cognitive and metacognitive strategies to students.

4.6.3.1.3. Texts

The texts used in the intervention for both strategy instruction and the control group were taken from magazines, newspapers, books and other sources to mimic real world reading tasks. Special attention was paid to ensure a diversity of familiar and unfamiliar topics. Each text was followed by a task which the students had to complete. The task could be to answer comprehension questions, provide a short summary, compare parts of the text with the students' lives, etc. These were designed to challenge the students' knowledge and comprehension, and give them the opportunity to apply the strategies they had learned.

As noted above, the students' comprehension was assessed after the first stage of the intervention using a version of the reading comprehension test. Also, the students answered the metacognitive awareness instrument in order to record whether their initial awareness of strategies and their self-reported use of strategies had changed.

4.6.3.2. Second stage: Immediate post-test and classroom teacher intervention

The study included the participation of schoolteachers to provide a more realistic context for the instruction. This characteristic followed the specifications of a pragmatic trial. A pragmatic trial sets a treatment in conditions as similar as possible to those present in the context where the treatment will be applied. This close match makes it possible to measure the effectiveness of the treatment in a realistic setting and improves the external validity (D. Torgerson and C. Torgerson, 2008).

The OECD reports high quality and regular in-service training delivered to teachers as a common characteristic of high-achieving countries in international tests (UNESCO, 2004). Also, research suggests that providing teachers with school-embedded professional learning opportunities instead of being trained during short-term workshops improves effectiveness (Dillon et al., 2011). Therefore, two schoolteachers were invited to participate in the project to enhance the external validity of the study. Incorporating teachers into the intervention provided a more realistic context for the instruction and contributed to teacher development. Both teachers spent the same amount of time delivering Spanish lessons[9] to the five classes. It was important to ensure that both teachers would have the same amount of contact time in order to avoid confounding due to differences in teachers' characteristics (L. Cohen, Manion and Morrison, 2011).

Effective teacher development can be delivered using literacy coaching models, which involve the teachers working cooperatively in either small groups or in pairs (Neufeld and Roper, 2003). Between the first and second stage of the study, there was a period in which the teachers were trained how to teach cognitive and metacognitive reading strategies as well as to practise with students the strategies already learnt. This period also helped the students in both groups to become accustomed to working with a new instructor. The training of the teachers was carried out over three weeks following a literacy coaching model (Neufeld and Roper, 2003). During the first week, the researcher met with the teachers to explain the differences in teaching approach and learning activities in the control and strategy group. They were explicitly instructed on how to practise the reading strategies, use modeling, provide explanations and give feedback. The importance of outlining the three types of knowledge was emphasised. After the initial meeting, the teachers observed the

[9] Teachers delivering the Spanish subject are in charge of teaching grammar, spelling and literature. They were invited to participate in the project because the activities were closely related to their subject and they would be interested in developing reading comprehension skills among their students.

instruction delivered by the researcher to both groups (strategy and control). After the lessons, the teachers and the researcher met again to discuss the lesson and address any further questions. In the second and third week, the teachers delivered the lessons and the researcher provided feedback on the teachers' performance. Thereafter, the teachers delivered the lessons independently.

The materials were designed by the researcher to ensure control of the difficulty and variety of materials. To monitor and control the teaching, the teachers were provided with a digital recorder and asked to record one lesson per week with each group. These recordings were sent on a weekly basis to the researcher. They were monitored by the researcher who provided support and further comments to the teachers, when necessary. The teachers were advised about how to elicit the strategies efficiently, how to help the students to express aloud the procedures followed to understand the text, and to encourage students to describe the rationale behind the use of chosen strategies. This support was provided during the two and a half months allocated to the strategy instruction delivered by the teachers. During this second stage, the students' worked independently - organising the work by themselves in small groups. The teachers worked with whole class (i.e. 35 students). In addition to teacher development and the ecological validity of the study, the practice of reading strategies during the second stage provided the students in the strategy group with the opportunity to practise the strategies by switching from semi-controlled activities to independent work. The teachers monitored the students' work as well as the organisation of the teams. They did not direct the work done.

At the end of the second phase, students took the reading comprehension and MAI post-test.

4.6.3.3. Third stage: Delayed post-test

At the end of the instruction delivered by the teachers, the students were encouraged to continue practising the strategies learnt at their own pace. For two months, they did not receive any further strategy instruction. The lessons followed the reading activities included in the syllabus and no extra reading practice or support was provided. After this time, the students' comprehension was reassessed using the third version of the reading comprehension test. This late assessment evaluated their comprehension after a period of teacher-delivered instruction (two months) and another period of no intervention (two months). Researchers have pointed out that short interventions are likely to yield positive results due to novelty and intensive

training (Pressley, 2000). The maintenance of the results, however, is an aspect rarely assessed in experimental studies. Also, linking back to learning theory (Mitchell, Myles and Marsden, 2013), the process of automatisation is very long - so it is very important to give lots of opportunities for practice.

In summary, the students' comprehension was assessed on three occasions: 1) at the beginning of the intervention, 2) directly after the instruction delivered by the researcher, and 3) two months after the end of instruction delivered by the teachers, i.e. four months after the end of the instruction by the researcher.

4.7. Statistical analysis

The statistical literature in the social sciences describes a range of methods for drawing inferences about different types of data. The most appropriate statistical approach to analysing data depends on the characteristics of the data collected. For continuous and approximately normally distributed data, independent sample t-test and analysis of (co)variance (ANOVA/ANCOVA) were used to assess group differences. If the assumption of normality was violated, this was indicated and a non-parametric alternative was performed instead. The data employed in all tests of continuous variables were tested for normality of distribution and homogeneity of variance. Normality of distribution was tested using the Kolmogorov-Smirnov test. In addition, the skewness ratio and kurtosis ratio were calculated (A. Field, 2009). The homogeneity of variance was tested using Levene's test. Where statistical procedures required additional assumptions (e.g. homogeneity of regression slopes in ANCOVA), these were assessed and reported in the relevant section.

In the case of the MAI, answers to each question were collected using 4-point rating scales. There is some controversy about how such data should be analysed. Some authors emphasise that this type of data is categorical[10] - the answers can be ordered, but the difference between them has no interval interpretation (A. Field, 2009; Agresti, 2007). Accordingly, appropriate statistical methods for ordered categorical data must be employed, such as Chi-square test, Fisher's exact test or ordered logistic regression (Agresti, 2007). However, many applied research studies in the field of education treat such data as interval (i.e. continuous) data and employ inferential

[10] Consider, for example, a question about the perceived effectiveness of reading strategies where the responses are collected using a 5-point Likert scale: 1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree. The responses can be ordered by the strength of agreement but it is unclear whether the move from 'disagree' to 'neutral' expresses the same increase in agreement than that from 'agree' to 'strongly agree' (Agresti, 2007).

statistics that are appropriate for interval data only. This approach implicitly assumes a similar difference between each level on the rating scale. Because of this controversy, it was decided to analyse the data obtained from the MAI using both approaches: treating them as interval data and as categorical data.

Whenever the results of the statistical tests are reported in this study, the significance of these were assessed using a critical value of $\alpha = .05$, which is commonly used in the social sciences. This implies that, if the experiment were to be repeated 100 times, we would expect the null hypothesis to be rejected incorrectly in five instances. Effects that were borderline significant ($0.05 < \alpha \leq 0.055$) are also highlighted.

4.8. Pilot study

In November 2010, a pilot study was conducted in the same school as the main study to test the instruments and the feasibility of the intervention. As in the main study, five classes participated in the pilot with a total of 171 students enrolled in their last year of secondary education. All classes participated in the piloting of the reading comprehension tests. The students from a class selected by the principal and the teachers participated in a five-hour strategy intervention. This class was chosen, taking into consideration the teachers' professional judgement - it was described as representative of the school population. The background questionnaire and the MAI were piloted with this class during a 50-minute morning lesson. The other four classes only took one of the three versions of the reading comprehension test. The design of the instruments is fully described in Section 4.5. The following sections provide a description of the activities in the short pilot intervention, an analysis of the instruments in the pilot stage of the study, and a report on the changes made as a result of the pilot study.

4.8.1. Pilot Word-level study

A subsample of 17 students participated in the pilot WLS. The assessment comprised of two sections: word identification and word comprehension.

In the word identification test, the students read aloud 35 Spanish words (see Section 4.5.1.1). In the word comprehension section, the students identified 20 synonyms, 20 antonyms, and 15 analogies. The sentences used for this test were taken from the three study guides mentioned before. On average, the students were able to identify 92.4% of the words correctly. The results of the word identification

test (mean=32.35, SD=3.02) suggest that the students did not encounter decoding problems. In contrast, the results of the word comprehension section (mean=22.12, SD=6.40, total possible = 55) indicated either poor vocabulary knowledge or that the students were unfamiliar with these specific words, which might indicate that the test was not a good indicator of lexical knowledge. No reference was found to support the validity of the study guides. Therefore, a different instrument was chosen to test vocabulary comprehension (WISC-R, as discussed in Section 4.5.1.2).

4.8.2. Pilot Metacognitive Assessment Instrument

MAI measured the students' perceptions of their strategy use. It was designed based on two self-reported instruments (see Section 4.5.3). An initial list of 50 items was piloted with 30 students. They were told to report the frequency of their strategy use on a five-point rating scale, then write down any problems they had with the statements and indicate if any of them were difficult to understand. Based on the students' responses, two items (Q16, Q17) were reworded and a four-point rating-scale was defined to measure each statement. Two items with very similar content (Q9, Q10) were merged. It was impossible to make comparisons with a post-test because the MAI was tested only once during the pilot.

4.8.3. Pilot strategy intervention

The students in the class chosen to participate in the pilot intervention took the reading comprehension test (version A). Base on their scores, the students were pair matched and then randomly allocated to one of two groups: the strategy group and control group. Both groups participated in the five-hour instruction distributed over eight 40-minute sessions. Both groups read the same texts and carried out the same designated activities.

The students in the strategy group received direct strategy instruction in two strategies: summarising and question elaboration. In the first session, the students in the strategy group discussed the importance of reading strategies, elaborated a list of useful strategies and were taught the steps needed to summarise a text (L. A. Brown and Day, 1983). In the next three sessions, the students worked on different texts to summarise the information. They worked in teams to practise the strategy and share ideas. During the following four sessions, the students worked on question elaboration following the same procedure. The participants in the non-intervention group read the same text and carried out activities such as underlining and discussing

unknown or difficult words, listening to the teacher's explanation of the topic and completing the exercises. No change was implemented to the intervention tasks as a result of the pilot study. The materials piloted formed a subset of those used in the main study.

4.8.4. Pilot reading comprehension test: test version equivalence

Three versions of the reading comprehension test were constructed (A, B, and C) to avoid improvement due to memorization of the items and to increase the range of content tested. Each version consisted of 5 texts with 3 or 4 items (i.e. multiple choice and open questions). Students from the remaining four classes were randomly allocated to take one of the three versions of the comprehension test (version A: N=107, version B: N=74, version C: N=116). They worked with the researcher twice: at the beginning of the pilot (pre-test) and at the end of the pilot (post-test). Two analyses were carried out to ensure equivalence of the three versions: item analysis and pre-test/post-test analysis. The item analysis was conducted across all tests to examine the suitability and equivalence of the items. The pre-test/post-test analysis compared the mean scores of the three versions of the test to examine the differences among the tests.

4.8.4.1. Item analysis

The item analysis sought to determine the item discriminability and item difficulty (L. Cohen, Manion and Morrison, 2011) of each comprehension item. It was conducted on the pre-test and post-test versions. That is, an item analysis was carried out for version A taken as a pre-test and another for the same version taken as a post-test.

Item discriminability indicates the effectiveness of the item to differentiate between groups of students. The comprehension test scores are ranked and divided into 3 groups: high, medium, and low scores. The score from the low scoring group (L) is subtracted from the score of the high scoring group (H) to calculate the difference. This number is then divided by half the number of students in the two groups:

$$\frac{H - L}{\frac{1}{2}N} \tag{4.1}$$

Item difficulty is calculated by dividing the number of students who answered the item correctly (Z) by the number of students who attempted the item (N). The result is multiplied by 100 to create an index of difficulty.

$$\frac{Z}{N} * 100 \tag{4.2}$$

L. Cohen, Manion and Morrison (2011) consider appropriate items to be those that fall between 30 and 67 per cent of the index [11]. Items Q2, Q6, Q7, Q8, Q18, and Q19 in the three versions fell outside this range. Therefore, they were reviewed and amended for the main study. The results of these analyses identified students who attempted to answer the question and achieved this fully or partially (i.e. one or two points). A second analysis was carried out for items Q3, Q8, and Q13 because a more complex process was required in order to answer these items correctly (i.e. reflection). Thus, few students obtained the full score (two points per item). This second analysis for these items only included data from students who obtained full marks for these items.

4.8.4.2. Pre-test and post-test analysis

The results from the descriptive statistics indicated a normal distribution of pre-test scores (all versions N=171) (mean=10.34, SD=3.66). Therefore, a one-way mixed analysis of variance (ANOVA) was conducted to test for differences between the three pre-test versions of the tests: A, B and C. The results indicated no statistically significant difference at the $p<.05$ level in the pre-test scores for the three versions ($F(2,169)=2.30$, $p=.103$). Therefore, the three versions were considered as equivalent. A separate analysis of the post-test scores was not performed because the results might have been affected by the pilot intervention.

The same analysis was conducted for all test versions regardless of the time of application (pre-test and post-test), i.e. the data were pooled but the post-test scores of the students in the intervention groups were excluded. Because the data did not meet the assumption of normality and homogeneity of variance, a Kruskal-Wallis test was carried out. This test did not reveal a statistically significant difference between the different versions of the test (version A: N=107, version B: N=74, version C: N=116), $\chi^2(2,297)=2.58$, $p=.275$.

In summary, the results from the item analysis, the ANOVA for the pre-test and post-test version and nonparametric tests, and the readability indexes of the texts indicated that there was a good level of equivalence between the tests. However, it was noticed that the test was too long for the allocated 50 minute time slot. Therefore,

[11] A higher value of the index indicates a more difficult item.

the last text, a non-continuous text advertising services, was eliminated from the three versions.

4.8.5. Pilot verbal reports

Practical problems beyond the researcher's control (i.e. unexpected school activities) impeded the pilot testing of the verbal reports.

4.8.6. Results of the pilot intervention

While the pilot study was primarily concerned with assessing the suitability of the materials, tests and instruments, an analysis of the (rather short) intervention was also performed. A one-way between-group ANOVA was conducted on the pre-test scores to test for differences at baseline between three experimental conditions: the intervention group (mean=9.25, SD=5.04), comparison group (mean=8.76, SD=4.15), and test only group (mean=10.84, SD=3.97) [12]. The results indicated that the mean scores did not differ statistically significantly between the groups (F(2,56)=1.28, p=.286). Therefore, the three groups were considered similar at baseline.

The same procedure was followed for the post-test mean scores. The results from the one-way ANOVA indicated that the means of the groups (intervention group: mean=11.68, SD=3.59; comparison group: mean=12.65, SD=2.87; and test-only group: mean=12.24, SD=3.85) did not differ significantly (F(2,48)=.35, p=.710). Therefore, it was concluded that no difference existed among the groups after the intervention.

Finally, a two-way mixed ANOVA was conducted to assess the impact of the three different conditions (intervention, comparison and test-only) on the participants' comprehension scores. There was no significant interaction between the conditions and the time of the application of the test (pre-test and post-test), Wilks' λ=.92, F(3,46)=2.09, p=.130. There was a substantial main effect for time, Wilks' λ=.65, F(3,46)=24.90, p=<.001, partial eta squared=.35, with the three groups showing an increase in post-test scores. The main effect comparing the three conditions was not significant, F(2,47)=.47,p=.628, partial eta squared=.02, suggesting that there was no difference among the three conditions.

As stated above, the main objective of the pilot was to test the equivalence of the comprehension tests, the logistics of the work within the school, the work style of the

[12]This only included scores from pre-test version B and post-test version C.

school, and the kind of data that would be elicited and its consequence for the data analysis. The small sample size and short intervention period were considered as potential explanatory factors for the absence of effect in the small pilot intervention. It should be kept in mind that the pilot study was not designed to estimate the effect of the intervention and the findings presented here should not be interpreted as evidence that teaching reading strategies is ineffective in increasing comprehension. The main study served to answer this question.

5. Results of the reading comprehension tests

The following two chapters report the results of the tests carried out to answer the five research questions posed for this project (see Chapter 3). This chapter covers the results of the quantitative analysis of the reading comprehension instruments that were used to answer the following research questions:

- Q1: Does the direct teaching of cognitive and metacognitive reading comprehension strategies help 15-year old Mexican students to improve their reading comprehension of Spanish texts?
- Q5: Do other factors, such as indicators of socio-economic status, baseline reading comprehension scores, and word recognition, influence the effectiveness of the intervention on reading comprehension?

Statistical analyses of the changes in reading comprehension scores after the intervention served to answer this question. To ensure that the poor comprehension was not due to poor decoding, the students' basic reading ability was assessed using the three WLS subtests (see Chapter 4). This chapter presents both the results of the WLS tests and those of the three reading comprehension tests (i.e. the pre-test, immediate post-test and delayed post-test).

It is important to note that the results in the present and following chapters are not always based on the entire sample of students (N=177). Some tests are applied separately to each of the two groups (i.e. strategy and control) and some tests are applied to a subsample of 62 students taken from the overall study population. To avoid confusion, here and thereafter, these groups will be referred to as: study sample (N=177) and subsample (N=62).

5.1. Word-level study

The WLS is composed of three subtests aimed at verifying the students' basic reading skills (see Section 4.5.1). A subsample of 62 students, randomly chosen from the

original study sample, took the tests: 19 from the control group (26% of the total N=74) and 43 from the strategy group (42% of the total N=103). For practical reasons, not every student took all three tests. However, this is unlikely to affect the interpretation of the results of the WLS seriously because the results of each test were analysed independently. Data from one visually impaired student in the strategy group were excluded from the analysis because his medical condition did not allow him to participate in the test. Hence, 61 students participated in the WLS study.

5.1.1. Word Identification test

5.1.1.1. Words grouped by frequency

Sixty-one students participated in the word identification test, of which 19 were members of the control group and 42 were members of the strategy group. On average, these students identified 95.3% of the words (mean=33.36 out of 35 words, standard deviation (SD)=2.02) (see Table 5.1), which indicates that they were able to decode words in the Spanish language. The results from an independent samples t-test showed no statistically significant differences between the number of words identified by the students in the control group (mean=33.42, SD=1.81) and the strategy group (mean=33.33, SD=2.13) (t(59)=.16, p=.88, r=.021). Data were also analysed using the Mann-Whitney test, a non-parametric alternative, due to negative skewness in the distribution of the data. The results showed that there were no differences between the groups (U=326, p=.65). This finding was robust to the exclusion of four outliers, i.e. observations with extreme values (above or below 2 SD from the mean).

The words employed in the test were grouped according to their frequency into high, medium, and low (see Appendix C.1) and the total score of identified words for each student was calculated. On average, the students correctly identified 98.8% of the high frequency words, 95.5% of the medium frequency words, and 93% of the low frequency words. Paris and Hamilton (2008) suggest a threshold of 90%, with students exceeding this threshold being unlikely to suffer from decoding problems, which could lead to comprehension problems. Therefore, the results were considered to indicate that the students in the study possessed sufficient word level decoding skills. The results from an independent sample t-test revealed no statistically significant differences between the strategy and control groups with regard to high frequency words (t(59)=.55, p=.58, r=.071), medium frequency words (t(59)=-.23, p=.82, r=.029), and low frequency words (t(59)=.22, p=.83,

r=.028).

Frequency	Group	Mean	SD	% of words
High	Control	7.95	0.23	99.3
(8 words)	Strategy	7.88	0.50	98.5
	Subsample	7.90	0.44	98.8
Medium	Control	12.37	0.83	95.1
(13 words)	Strategy	12.43	0.99	95.6
	Subsample	12.41	0.94	95.5
Low	Control	13.11	1.20	93.2
(14 words)	Strategy	13.02	1.39	93.6
	Subsample	13.05	1.32	93.0
Total	Control	33.42	1.81	95.5
(35 words)	Strategy	33.33	2.13	95.2
	Subsample	33.36	2.02	95.3

Table 5.1.: Word identification grouped by frequency

5.1.1.2. Words grouped by number of syllables

The data were also grouped by the number of syllables using the following classification: 2-3, 4-5 and 6 or more syllables. The scores for identified words were calculated to test whether this form of grouping words would yield different results to the frequency grouping. The results presented in Table 5.2 show that the students in both groups identified 96.5% of words with 2-3 syllables, 93.4% of words with 4-5 syllables, and 96.7% of words with 6 or more syllables. The number of correctly-identified words was not normally distributed and the non-parametric Mann-Whitney test was employed. The results of this test showed that the students in the control group's ability to identify words (mean=33.42) did not differ statistically from that of those in the strategy group (mean=33.33) with regard to 2-3 syllable-words (U=333, p=.215, r=-.16), 4-5 syllable-words (U=339, p=.313, r=-.13), and 6 or more syllable-words (U=387, p=.780, r=-.036) [1].

[1]The effect size was calculated using the formula $r = \frac{Z}{\sqrt{N}}$ taken from A. Field and Hole (2003, p. 238).

Frequency	Group	Mean	SD	% of words
2-3 syllables	Control	13.68	0.67	97.7
(14 words)	Strategy	13.43	0.86	95.9
	Subsample	13.51	0.81	96.5
4-5 syllables	Control	12.95	1.13	92.5
(14 words)	Strategy	13.14	1.26	93.9
	Subsample	13.08	1.22	93.4
$\geq$6 syllables	Control	6.79	0.54	97.0
(7 words)	Strategy	6.76	0.53	96.6
	Subsample	6.77	0.53	96.7
Total	Control	33.42	1.81	95.5
(35 words)	Strategy	33.33	2.13	95.2
	Subsample	33.36	2.02	95.3

Table 5.2.: Word identification grouped by number of syllables

5.1.2. Word knowledge test

The word knowledge tests consisted of two components: the vocabulary and the word similarities test. Each test was taken by 47 students: 17 out of the 74 students (23%) in the control group and 30 out of the 103 students (29%) in the strategy group. The highest possible score was 64 points for the vocabulary test, and 30 points for the word similarities test. Data were normally distributed so that any differences between the groups were assessed using independent sample t-tests.

The results from the vocabulary test showed that the students in the strategy group possessed a better vocabulary knowledge (mean=51.17, SD=6.00) than the students in the control group (mean=46.47, SD=8.60), and this difference was statistically significant (t(45)=-2.21, p=.032, r=.31). The implications of this finding will be discussed in Chapter 7. In the word similarities test, the results indicated a better performance by the students in the strategy group (mean=23.53, SD=3.94) compared to those in the control group (mean=22.71, SD=3.80). This difference, however, was not statistically significant (t(45)=-.702, p=.486, r=.10).

The results of the vocabulary test were analysed to assess whether a relationship exists between the frequency of the words included in the vocabulary test (taken from

Davies (2002)) and the students' ability to define the word. The Pearson's correlation coefficient (ρ) was not statistically significantly different from zero (ρ=-.11, p=.535), suggesting that there was no relationship between the frequency of the word and the number of times the word was identified.

The total scores of the word knowledge tests were also compared against the assessment scales provided by the WISC-R test. These scales are used to identify the appropriate level of students' vocabulary knowledge according to their age. A comparison against the WISC-R scales was made to see whether the students in the subsample were comparable to the general population of students of the same age. The scales show the expected vocabulary knowledge at different ages by placing students within 19 performance levels. Students possessing excellent vocabulary knowledge are placed on the 19th level and students with poor vocabulary knowledge are placed on the 1st level. Table 5.3 provides an example of the WISC-R scale for students aged 15 and 4 to 7 months.

The frequencies of vocabulary test placed the majority of students in both groups within levels 11 (19%) and 12 (19%), while the frequencies of the word similarities test place them in levels 13 (15%) and 16 (17%) (Tables 5.4 and 5.5). These results were interpreted as evidence that the students performed adequately on both tasks given their age, and confirmed the findings of the previous section (Section 5.1.2).

WISC-R levels	Similarities score	Vocabulary score
1	0-6	0-19
2	7	20-21
3	8	22-23
4	9-10	24-26
5	11-12	27-29
6	13-14	30-32
7	15	33-36
8	16-17	37-39
9	18-19	40-43
10	20	44-46
11	21	47-49
12	22	50-52
13	23	53-54
14	24	55-56
15	25	57-58
16	26	59
17	27	60
18	28	61
19	29-30	62-64

Table 5.3.: Example of WISC-R scale for students aged (years/months) 15/04 to 15/07

Levels	7	8	9	10	11	12	13	14	15	16	17	18	19
Control group (N=17)													
Count	1	1	2	1	5	2	1	0	1	1	1	1	0
Percentage	5.9%	5.9%	11.8%	5.9%	29.4%	11.8%	5.9%	0.0%	5.9%	5.9%	5.9%	5.9%	0.0%
Strategy group (N=30)													
Count	0	1	0	1	4	7	5	3	5	1	1	0	2
Percentage	0.0%	3.3%	0.0%	3.3%	13.3%	23.3%	16.7%	10.0%	16.7%	3.3%	3.3%	0.0%	6.7%
Overall (N=47)													
Count	1	2	2	2	9	9	6	3	6	2	2	1	2
Percentage	2.1%	4.3%	4.3%	4.3%	19.1%	19.1%	12.8%	6.4%	12.8%	4.3%	4.3%	2.1%	4.3%

Table 5.4.: Classification of students according to vocabulary WISC-R scale

Levels	7	8	9	10	11	12	13	14	15	16	17	18	19
Control group (N=17)													
Count	0	3	0	0	0	3	4	3	0	1	1	2	0
Percentage	0.0%	17.6%	0.0%	0.0%	0.0%	17.6%	23.5%	17.6%	0.0%	5.9%	5.9%	11.8%	0.0%
Strategy group (N=30)													
Count	1	1	1	1	3	2	3	3	3	7	0	4	1
Percentage	3.3%	3.3%	3.3%	3.3%	10.0%	6.7%	10.0%	10.0%	10.0%	23.3%	0.0%	13.3%	3.3%
Overall (N=47)													
Count	1	4	1	1	3	5	7	6	3	8	1	6	1
Percentage	2.1%	8.5%	2.1%	2.1%	6.4%	10.6%	14.9%	12.8%	6.4%	17.0%	2.1%	12.8%	2.1%

Table 5.5.: Classification of students according to similarities WISC-R scale

Because the classification results are ordinal rather than interval data, a Fisher's exact test was used to compare the groups (see Section 4.7 for a discussion of this approach) [2]. The results from the vocabulary scales in the WISC-R test showed that the median student in the strategy group was on a higher WISC-R level (level 13) than the median student in the control group (level 11). However, the distribution of scores in the control group did not differ from those in the strategy group ($p = .209$). In the word similarities test, the results indicated a slightly better level for students in the strategy group (level 14) than for students in the control group (level 13) but, again, the distribution of scores in the control group did not differ from those in the strategy group ($p = .232$). These results confirmed that, for the vocabulary and similarities test measured by the WISC-R scales, there was no statistically significant difference between the groups. The results were also interpreted as evidence that the students performed adequately on both tasks, given their age, and confirmed the findings of the previous section (Section 5.1.2).

As before, it was found that the students performed better on the word similarities test than on the vocabulary test. Hence, it might be concluded that the students in the sample possessed a better understanding of vocabulary when words were paired and/or could be used within a context than when they were asked to provide a definition of them.

5.1.3. Summary of lessons arising from the word level study

The results from the WLS indicated that the students in the sample were adequately skilled in identifying Spanish words and that they possessed good vocabulary knowledge. They were also placed within the high performance levels (11 and above) on the WISC-R scale. However, the results indicated a worse performance in accessing meaning when describing isolated words (i.e. in the vocabulary test) than when accessing word meaning within a context (i.e. finding similarities between words). Several reasons might explain the students' performance in the vocabulary test. First, the students' verbal ability might affect the results. It might be hypothesised that the students were able to recognise the word (as suggested by the results of the similarities test) but were unable to define it accurately due to poor verbal fluency. They might also have been unaccustomed to performing this kind of task. This point was not further investigated as it is beyond the scope of the study. Another reason for

[2] Fisher's exact test was used because the data did not follow the assumption of the Chi-square test, i.e. some cells had an expected frequency of five or less. Fisher's exact test requires no such assumption.

the lower scores for the vocabulary test might be the strict marking criteria used in the test; points were only awarded when correct descriptions were provided. When explaining adjectives and nouns, some students either used a different part of speech coming from the same lexical family (lexeme), such as a verb (i.e. contagioso - contagiar, contagious - infect; espionaje - espiar; espionage - spy), or provided sentences containing the word to be described (i.e. apostar en el juego, tener rivalidad con alguien). The students were told not to use the same words within a sentence or words from the same lexical family to explain the words. Even so, it was observed that they frequently used lexemes and hence were not granted full marks.

Figure 5.1 shows the percentage of words that the students in the strategy and control groups could identify and define correctly during the three WLS tests. In the word identification test, the achievement of both the strategy and control groups exceeded 90%. However, in the vocabulary and word similarities test, the students could only correctly define approximately 70-80% of the words. These results are similar to findings reported by other authors (see Chapter 3 and Madariaga Orbea and Martínez Villabeitia (2010)) who concluded that poor vocabulary knowledge might be one of the factors leading to reading comprehension deficiencies. The results show that there was no statistically significant difference between the strategy and the control group for the word identification and word similarities tests. In the vocabulary test, the strategy group showed statistically significant better vocabulary knowledge than the control group when a comparison was made using a points scoring scale but not when the comparison was made using the WISC-R scales.

5.2. Reading Comprehension Test

The results from the three versions of the reading comprehension test mainly served to answer the first research question (Q1). This section also compares between groups the students' reading comprehension at baseline (pre-test). The acknowledgement of any differences between the groups before the intervention is important because true randomisation at the individual level was impossible due to practical reasons (see Chapter 4). These issues are discussed further in Chapters 7 and 8.

5.2.1. Exclusion criteria

Data were collected from the original sample population of 177 students. However, this number was reduced throughout the study for three reasons: firstly, 7 students

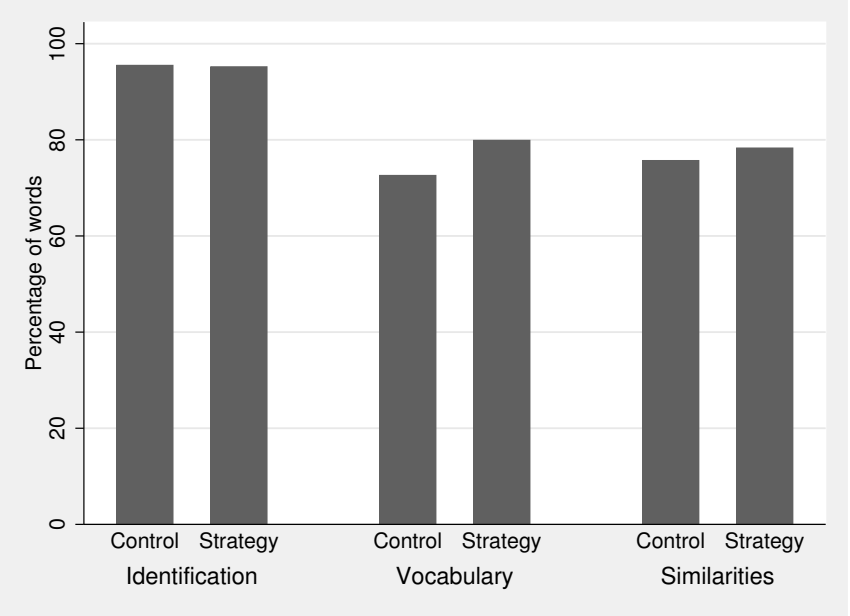


Figure 5.1.: Percentage of words correctly identified and defined

were excluded from the study due to different reasons (e.g. students changing school, being expelled from school, or having been diagnosed with sight or learning disabilities) (Appendix D.1); secondly, 27 students who participated in the intervention but failed to take one or more tests were excluded; and, finally, three students were excluded whose pre-test, immediate post-test, or delayed post-test scores were above or below 3 standard deviations from the mean [3]. Table 5.6 shows the number of participants in each group who were excluded for each reason and the total number of participants included in the analysis.

5.2.2. Check of normality assumptions

Descriptive statistics for the pre-test, post-test and delayed post-test scores were calculated for the entire sample and for each group separately to check whether the data were normally distributed. Data for the entire sample were approximately normally distributed in the delayed post-test but not in the pre-test and immediate post-test. However, A. Field (2009) emphasises that the more important question is whether the distribution of data for each group is approximately normal. The results from the Kolmogorov-Smirnov test indicated that the data for the control group were

[3] This was done to guard against the influence of extreme observations in any of the three tests.

Group	Initial Sample	Reason 1: Dropout	Reason 2: Missing test	Reason 3: Score above or below 3 SD	Analysis sample
Control	74	2	11	0	61
Strategy	103	5	16	3	79
Study sample	177	7	27	3	140

Table 5.6.: Exclusion criteria for reading comprehension study

not normally distributed in the pre-test and that the data for the strategy group were not normally distributed in the immediate post-test. However, the calculated skewness and kurtosis ratio were below the critical z-score value of 2, suggested by A. Field (2009), for either group. This indicated that it was appropriate to consider the data for each group to be normally distributed for the purposes of statistical analyses. In addition, Weinberg and Abramowitz (2002) appeal to the Central Limit Theorem, stating that *"sample sizes larger than 30 eliminate the need to check data for normality"* (Weinberg and Abramowitz, 2002, p. 276). The sample sizes for both groups were above this threshold.

5.2.3. Groups at baseline

The allocation of participants to the control and strategy groups (see Chapter 4) did not follow a randomised process at the level of the individual. Thus, it was important to measure the students' level of reading comprehension before the start of the intervention and take any pre-existing differences between the groups into account during the statistical analysis. Understanding the characteristics of the sample allowed a better acknowledgement of the differences between the groups at baseline and the interpretation of the results (Moore, A. Graham and Diamond, 2003; Shadish, Cook and Campbell, 2002).

Descriptive statistics for the pre-test scores showed that both groups met the normality assumption. The results from an independent sample t-test indicated that, at baseline, the control group (mean= 11.82 out of 21, SD = 2.82) had a higher reading comprehension than the strategy group (mean = 10.97, SD = 2.94), but this difference was not statistically significant ($t(138) = 1.72$, $p = .085$, $r = .144$).

The effect size estimate indicated that the difference between the groups at baseline represents a small to medium effect. The statistical power of the test was 73% (J. Cohen, 1988), so the results were considered trustworthy.

5.2.4. Analysis of variance

The students' reading comprehension was reassessed on two more occasions to find out whether the groups performed differently immediately after the intervention (i.e. immediate post-test) and four months after it ended (i.e. delayed post-test). The latter time point was used to test whether any change due to the intervention was maintained over time. An analysis of the immediate post-test and delayed post-test scores showed homogeneity of variance for both groups for both tests but a slightly skewed distribution for the immediate post-test scores. The skewness ratio was calculated and the results ($z < 2$) did not indicate a problem with the distribution (A. Field, 2009). Also, as stated before, the number of observations was sufficiently large to appeal to the Central Limit Theorem (Weinberg and Abramowitz, 2002). Descriptive statistics of the test scores by group are presented in Table 5.7. The behaviour of both groups is also visualised using boxplots (see Figure 5.2), which show that reading comprehension in both groups improved immediately after the intervention and decreased again after four months.

Measurement	Group	N	Mean	SD
Pre-test	Control	61	11.82	2.82
	Strategy	79	10.97	2.94
	Study sample	140	11.34	2.91
Immediate Post-test	Control	61	13.03	2.99
	Strategy	79	12.49	3.70
	Study sample	140	12.73	3.41
Delayed Post-test	Control	61	11.90	3.26
	Strategy	79	12.25	3.47
	Study sample	140	12.10	3.37

Table 5.7.: Descriptive statistics for groups across time

A two-way mixed ANOVA was carried out in order to test whether the differences

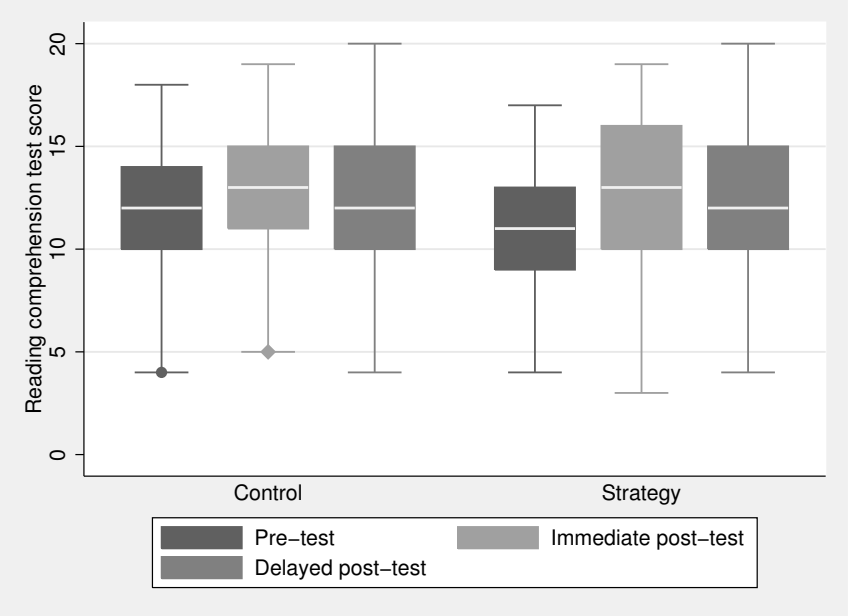


Figure 5.2.: Boxplots of reading comprehension scores

found in the descriptive statistics were significant and whether time and group membership had an effect on the students' reading comprehension. Mauchly's test indicated that the assumption of sphericity was not violated, $\chi^2(2)=.997$, p=.812. The results for the mixed ANOVA yielded a significant main effect of time on the students' scores, F(2,276)=12.82, p<.001, partial eta-squared=.085, power=.99 [4]. This indicated that the students' reading comprehension varied across time. The main effect of group was non-significant, F(1,138)=.585, p=.446, partial eta-squared=.004, power=.12. This indicated that, regardless of when reading comprehension was measured, the students' reading comprehension in the control group was not significantly different from that in the strategy group. The interaction between time and group was non-significant, F(2,278)=2.53, p=.081, partial eta-squared=.018, power=.50.

To gain a better understanding of the differences in time, it was important further to explore the effect of time by carrying out paired comparisons. As seen in Table 5.8, the results from the posthoc tests showed an initial significant improvement in both groups after the intervention (i.e. immediate post-test). However, there was a statistically significant drop in comprehension after four months (i.e. delayed

[4]The partial eta-squared is reported as effect size (Larson-Hall, 2010) because the sample size in both groups is different. A. Field (2009) suggests a formula to calculate *r* but this assumes the same number of participants in both groups.

post-test) in the control group (-1.13 points, p=.007, d=.46) and a smaller but non-significant drop in comprehension in the strategy group (-0.24 points, p=.505, d=.07). These findings indicate that, after the initial improvement of comprehension in both groups, the students in the strategy group maintained the effects over time (Figure 5.3). It is worth noting that no further explanations nor practice of strategies were carried out between the the end of teachers' instruction and the delayed post-test in any of the groups.

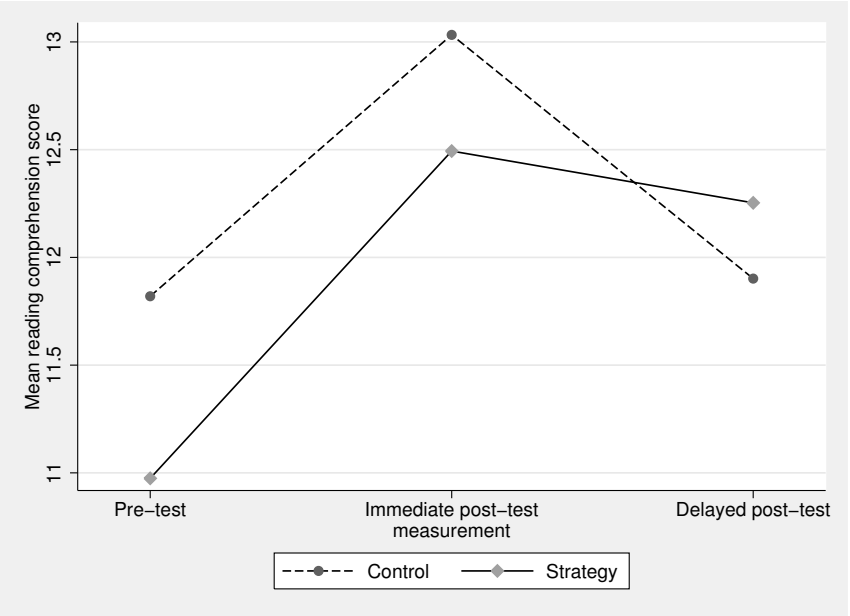


Figure 5.3.: Line chart of the groups means of reading comprehension scores over time

The results also showed an improvement in both groups when comparing the students' initial measurement of reading comprehension (i.e. pre-test) and measurement after four months (i.e. delayed post-test). Whilst the improvement among the strategy group was statistically significant (1.28 points, p=.001, d=.50), that among the control group was not (0.08 points, p=.848, d=.03). These findings indicate that the effect of the intervention was not maintained over time in the control group whilst the effect was maintained in the strategy group.

Several different factors other than the effect of the intervention itself may explain these results. First, the comprehension improvement in both groups immediately after the intervention (immediate post-test) might be due to the novelty of having an external person deliver a new workshop which the students might have considered

Group	Time A	Time B	Difference (B-A)	p-value
Control	pre-test	immediate post-test	1.21	.004
(N=61)	immediate post-test	delayed post-test	-1.13	.007
	pre-test	delayed post-test	0.08	.848
Strategy	pre-test	immediate post-test	1.52	<.001
(N=79)	immediate post-test	delayed post-test	-0.24	.505
	pre-test	delayed post-test	1.28	.001

Table 5.8.: Paired comparisons of mean reading comprehension scores

as beneficial for them (L. Cohen, Manion and Morrison, 2011). Second, there were increases in the classroom-based reading time during the intervention in both groups, which provided students in both the strategy and control group with extra practice. Third, there may have been a test-retest effect, i.e. students performing a similar test multiple times within a short time interval will improve due to familiarity with the test, not genuine improvements. However, this effect is unlikely to explain the observed effect because: a) tests were specifically designed to avoid this test-retest effect (see Chapter 4) and b) no such behaviour is evident in the delayed post-test scores. Finally, the control group was also exposed to some form of intervention ('active control group') as they were exposed to more reading than usual and were prompted more often to discuss texts. All of these points might help to explain part of the observed gains/losses. However, both groups faced the same threats to validity. The first three points (novelty of teacher variable, effect of reading material, and retest effect) were controlled for during the design of the project by ensuring that all elements were identical for both groups (i.e. teaching time, same texts and test questions, same instructor, etc.). The last point will be discussed further in Chapter 7.

Although the results of the mixed ANOVA indicated that the change in comprehension scores between the groups was not statistically significantly different, this finding might have been affected by the difference between groups before the intervention due to a lack of true randomisation. The difference at baseline of approximately one point (approx. 5% on overall scale), although not statistically significant, may have acted as a confounding variable which may have biased the results of the interaction. In addition, external factors such as the extra time students read for pleasure, socio-economic status, number of books at home, and attendance to the intervention might have an effect on the outcome. In order to control for the effect

of the differences between the groups at baseline and other confounding factors, two additional ANCOVAs were carried out as sensitivity analyses: one using the pre-test scores as a covariate, and the other using the pre-test and other confounding factors as covariates.

5.2.5. Sensitivity analysis: differences at baseline between the confounding factors

5.2.5.1. Analysis of covariance with pre-test scores as the covariate

The first ANCOVA included only the baseline (i.e. pre-test) scores and indicator variables for group membership (i.e. control and strategy) and time (immediate post-test and delayed post-test). The estimated coefficient β indicates the effect of the respective variable on the outcome when all other variables are held constant. The results indicated that the pre-test scores significantly predicted the students' scores for the immediate and delayed post-test, F(1,137)=54.53, p<.001, partial eta squared=.29. A one-unit increase in pre-test scores is associated with .587 higher scores in the immediate post-test (t-statistic=6.70, p<.001, partial eta squared=.25) and .519 higher scores in the delayed post-test (t-statistic=5.80, p<.001, partial eta squared=.20)(see Table 5.9). After controlling for this covariate, further paired comparisons of the groups revealed, as before, no significant differences between the groups for the immediate and delayed post-test, F(1,137)=0.73, p=.395, partial eta-squared=.01. Being in the control group rather than the strategy group was associated with .043 higher scores at immediate post-test (t=.09, p=.933, partial eta squared=0) and .790 lower scores at delayed post-test (t=-1.51, p=.133, partial eta squared=0.16).

	Immediate post-test				Delayed post-test			
Variable	β	SE	p-value	η^2	β	SE	p-value	η^2
Constant	6.06	1.02	<.001	0.21	6.56	1.04	<.001	0.23
pre-test	0.59	0.09	<.001	0.25	0.52	0.09	<.001	0.20
*Strategy group**	-	-	-		-	-	-	
Control group	0.04	0.51	0.933	0.00	-0.79	0.52	0.133	0.02

(*) Base category; β = beta; SE = standard error; η^2 = partial eta squared

Table 5.9.: Results of ANCOVA with pre-test scores as covariate

Comparisons over time revealed that, although the students in the strategy group's scores decreased slightly from the immediate post-test to the delayed post-test, the loss was not statistically significant (difference = -0.265 units, p=.465). Students in the control group, however, experienced a statistically significant loss in measured reading comprehension at four months after the intervention (difference = -1.099 units, p=.009). As expected, these results were in line with those from the Mixed ANOVA analysis reported above.

5.2.5.2. Further confounding variables

Previous research has shown that multiple factors affect reading comprehension (see Chapter 4). Four factors related to the students were tested in this study: number of books at home, the amount of time spent reading for pleasure, socio-economic status (SES), and attendance (% of lessons attended during the intervention). The first two were directly measured in the background questionnaire. The third factor was calculated as a composite measure of parental occupation and a family wealth score. The family wealth score was calculated based on the students' answers to the availability of certain objects in their house (e.g. a PC, an encyclopaedia, a study room, etc.). The highest values indicated more wealth. The students also provided information about their parents' occupation on the background questionnaire. This information was coded into nine categories according to the National Classification of Occupations in Mexico 2011 (Instituto Nacional de Estadística y Geografía, 2011), which is related to SES. The parent with the highest ranked occupation was chosen for each student. The score for parental occupation was added to that for family wealth to create a continuous SES score for the study. Finally, attendance was recorded by the researcher during the lessons.

Table 5.10 shows the frequencies for the number of books that the students report to have at home (grouped into four categories). Table 5.11 shows the frequencies for time spent reading for pleasure (grouped into five categories). These variables were compared between groups using Chi-square tests in order to test whether there were any differences at baseline.

The results of a Chi-square test did not reveal statistically significant differences between the groups regarding the number of books that the students possess at home (χ^2=3.44, p=.329), nor in the time that they spent reading for pleasure (χ^2=7.86, p=.097).

The average SES and attendance rates across groups were compared using in-

Number of books	Control group (N=61)		Strategy group (N=79)	
	Frequency	%	Frequency	%
≤10	12	19.7	20	25.3
11-25	20	32.8	17	21.5
26 - 100	18	29.5	31	39.2
>100	11	18.0	11	13.9

Table 5.10.: Number of books at home by group

Reading time	Control group (N=61)		Strategy group (N=79)	
	Frequency	%	Frequency	%
None	12	19.7	24	30.4
<30 min	30	49.2	25	31.6
30-60 min	7	11.5	18	22.8
61-120 min	9	14.8	7	8.9
>120 min	3	4.9	5	6.3

Table 5.11.: Reading time for leasure by group

dependent sample t-tests. The results showed no statistically significant difference between the SES scores of the control group (mean=16.74, SD=3.89) and the strategy group (mean=16.65, SD=3.98) (t(138)=.18, p=.891, r=.011), and no statistically significant difference between the attendance of the control group (mean=93.59, SD=10.09) and the strategy group (mean=93.58, SD=11.17) (t(138)=.004, p=.997, r=.0003).

In summary, there were no statistically significant differences between the control and strategy groups for any of the three variables that were collected as part of the background questionnaire and the attendance rate.

5.2.5.3. Analysis of covariance with pre-test scores and further confounding factors as covariates

The second ANCOVA included the same variables as the previous one (see Section 5.2.5.1) plus the four variables for the number of books at home, reading time, SES, and attendance rate. SES and attendance rate were included as continuous

variables and the other two were coded as categorical variables (see previous section). This ANCOVA was carried out to address two questions:

1. Are the four variables (number of books at home, reading time, SES, and attendance) statistically significantly associated with variation in the immediate and delayed post-test reading comprehension scores after accounting for the group assignment and pre-test scores?

2. Does the observed association between treatment assignment and immediate and delayed post-test scores, conditional on pre-test scores, change once these four potentially confounding factors are controlled for, i.e. are the findings of the previous ANCOVA robust in controlling for these factors?

The results indicated that the pre-test scores significantly predicted the students' scores for the immediate and delayed post-tests, $F(1,128)=47.70$, $p<.001$, partial eta squared=.27. A one-unit increase in pre-test scores is associated with .551 higher scores for the immediate post-test (t-statistic=6.27, $p<.001$, partial eta squared=.24) and .481 higher scores in the delayed post-test (t-statistic=5.24, $p<.001$, partial eta squared=.18)(see Table 5.12). Also, reading for pleasure for more than two hours, relative to not reading at all, significantly predicted the students' scores in the immediate and delayed post-test, $F(1,128)=5.45$, $p=.021$, partial eta squared=.04. Reading for more than two hours for pleasure is associated with 2.36 higher scores in the immediate post-test (t-statistic=1.98, $p=.050$, partial eta squared=.03) and 2.38 higher scores in the delayed post-test (t-statistic=1.94, $p=.054$, partial eta squared=.03) compared to not reading for pleasure (see Table 5.12). A one unit higher SES is associated with .152 higher scores in the immediate post-test (t-statistic=2.10, $p=.038$, partial eta squared=.03) but not in the delayed post-test. Attendance rates and number of books at home were not statistically significantly associated with reading scores.

After controlling for these four possibly confounding factors, further paired comparisons of the groups revealed, as before, no significant differences between the groups for the immediate and delayed post-test, $F(1, 128) = 0.174$, $p = .677$, partial eta-squared = .001. Being in the control group was associated with .297 higher scores at immediate post-test (t-statistic=.057, $p=.571$, partial eta squared=.003) and .669 lower scores at delayed post-test (t-statistic=-1.25, $p=.215$, partial eta squared=0.01).

	Immediate post-test				Delayed post-test			
Variable	β	SE	p-value	η^2	β	SE	p-value	η^2
Constant	3.07	2.65	.249	0.01	4.06	2.71	.137	0.02
pre-test	0.55	0.09	<.001	0.23	0.48	0.09	<.001	0.18
*Strategy group**	-	-	-	-	-	-	-	-
Control group	0.30	0.52	.571	0.00	-0.67	0.54	.215	0.01
Attendance	0.01	0.02	.636	0.00	0.02	0.03	.451	0.00
Socio-economic status	0.15	0.07	.038	0.03	0.02	0.07	.827	0.00
*Reading 0 min**	-	-	-		-	-	-	-
Reading <30 min	0.20	0.65	.752	0.00	-0.09	0.66	.894	0.00
Reading 30 - 60 min	1.38	0.79	.083	0.02	1.19	0.81	.143	0.02
Reading 61 - 120 min	0.83	0.95	.383	0.01	1.64	0.98	.097	0.02
Reading >120 min	2.36	1.19	.050	0.03	2.38	1.22	.054	0.03
*Books ≤10**	-	-	-	-	-	-	-	-
Books 11 - 25	-1.29	0.73	.082	0.02	0.31	0.75	.677	0.00
Books 26 - 100	-0.79	0.75	.298	0.01	0.60	0.77	.442	0.01
Books >100	-1.58	0.94	.095	0.02	0.46	0.96	.635	0.00

(*) Base category; β = beta; SE = standard error; η^2 = partial eta squared

Table 5.12.: Results of ANCOVA with pre-test scores and further potential confouding variables as covariates

Comparisons over time revealed that, although the students' scores in the strategy group decreased slightly from the immediate post-test to the delayed post-test, the loss was not statistically significant (difference = -0.208 units, p=.575). Students in the control group, however, experienced a statistically significant loss in measured reading comprehension at four months post-intervention (difference = -1.174 units, p=.006).

Overall, the effects of the ANCOVA with or without controlling for potentially confounding factors did not differ qualitatively from the results of the ANOVA. This indicates that, even though the students could not be individually randomised to the strategy and control groups, this is unlikely to have affected the results.

5.2.6. Summary of the results for the reading comprehension tests

The previous analyses help us to answer the question of whether the direct teaching of cognitive and metacognitive reading comprehension strategies help 15-year old Mexican students to improve their reading comprehension of Spanish texts. The results indicate that the direct teaching of reading strategies improved the students' reading comprehension and that, unlike the students in the control group's compre-

hension, the gains in the strategy group were maintained after a 4-month period without further strategy teaching.

6. Metacognitive questionnaire and verbal reports

This chapter presents the second part of the results of the project. It is divided into two parts. The first part describes the analyses of the data collected using the metacognitive assessment instrument (MAI) and the verbal reports. The results of the analyses served to answer the following research questions:

- Q2: Does the intervention raise the students' self-reported awareness of reading strategies and the reported frequency of use?
- Q3: Does the intervention increase the reported 'online' use of reading comprehension strategies when reading unfamiliar texts?

The analyses presented in the second part of the chapter combined the results of the different outcome instruments in order to gain further insights into the likely mechanism by which students in the strategy group improved and maintained their reading comprehension post-intervention. This served to answer the fourth research question:

- Q4: Is there a relationship between the self-reported use of strategies and the students' performance on the reading comprehension test?

6.1. Metacognitive questionnaire

The metacognitive questionnaire (MAI) consisted of 48 items, each of which required an answer on a four-point rating scale. The students were asked to indicate how often they used the reading strategy specified in each statement. The same version of the questionnaire was used before the intervention (MAI pre-test) and after the intervention (MAI post-test). The responses to these questionnaires were compared to test whether the students' perception of the use of strategies had changed over time. Comparisons across the groups allowed us to test whether any changes over time

were the result of explicit instruction in reading comprehension strategies. Finally, the analyses also served to identify which of the strategies the students reported using.

The next section is structured as follows: first, the exclusion criteria are explained. Second, an analysis based on aggregate scores with interval properties is presented. Third, the analysis is repeated and responses are treated as categorical data. Finally, a linear regression is performed to establish whether the results are explained by a set of potentially confounding variables (see Section 5.2.5).

6.1.1. Exclusion criteria

Data were collected from the original sample population of 177 students. The same exclusion criteria as in Chapter 5 were applied: firstly, 7 students were excluded from the study due to a range of different reasons, discussed previously (see Chapter 5); secondly, 4 students who participated in the intervention but failed to take either the pre-test (n=2) or the post-test (n=2) were excluded; and, thirdly, 2 students were excluded whose pre-test or post-test scores (interval scores; see below) were above or below 3 standard deviations from the mean. Table 6.1 shows the number of participants in each group who were excluded for each of these three reasons and the total number of participants included in the analysis.

Group	Initial Sample	Reason 1: Dropout	Reason 2: Missing test	Reason 3: Score above or below 3 SD	Analysis sample
Control	74	2	1	1	70
Strategy	103	5	3	1	94
Study sample	177	7	4	2	164

Table 6.1.: Exclusion criteria for metacognitive assessment instrument (MAI) study

Due to the controversy found in the literature regarding the best way to analyse self-report instruments, the responses were analysed using both methods: data were treated as interval and as categorical and the findings from each analysis were compared.

6.1.2. Responses treated as interval data

The answers to all items were scored from one to four and an overall score of the reported use of strategies was obtained by summing up across items. This resulted in two scores per student: one for the MAI pre-test and the other for the MAI post-test. The maximum score was 192 points for each test, with larger values indicating higher self-reported use of strategies. For the purpose of this analysis, the data were treated as interval data.

6.1.2.1. Comparison based on aggregate scores

6.1.2.1.1. At baseline

The scores for the MAI pre-test questionnaire were calculated and the results of the control (mean=117.83, SD=17.40) and strategy (mean=117.26, SD=17.00) groups were compared using an independent samples t-test. On average, the results showed no significant difference between the groups in the MAI pre-test based on comparisons of the interval scores (t(162) = .21, p=.833, r= .016).

6.1.2.1.2. At immediate post-test

The score of reported strategy-use during the MAI post-test served to test whether the students in both groups reported a different use of strategies after the intervention. The means for each group after the intervention were compared using an independent sample t-test. The results showed a statistically significant difference between the groups' perceived use of strategies. The students in the strategy group (mean= 130.29, SD = 18.44) reported a higher use than those in the control group (mean= 121.77, SD = 22.09)(t(162) = -2.69, p =.008, r = .206).

6.1.2.1.3. Gain scores

Gain scores were calculated by subtracting the MAI pre-test scores from the MAI post-test scores. The gain scores for both groups were compared using an independent sample t-test. On average, the gain scores of the participants in the strategy group were higher (mean = 13.03, SD = 16.90) than those of the participants in the control group (mean = 3.94, SD = 16.98). The difference between the groups was statistically significant (t(162) = -3.40, p = .001; r = .258).

In conclusion, these results indicated that the students' self-reported use of strategies was similar in both groups before the intervention. However, after the intervention, the students in the strategy group reported a more frequent use of strategies. Note that this analysis assumed interval properties of the aggregate scores. The results should therefore be treated with caution due to the reasons stated in Section 4.7 regarding the implicit interpretations of the scale of the responses.

6.1.2.2. Comparison based on scores clustered by strategy

The questionnaire items were clustered into reading strategies to identify which strategies the students reported using more often after the intervention. This analysis helped to test whether the direct instruction of reading strategies had changed the students' perceptions of strategy use and, if so, which strategies were reported to have been adopted. The analysis presented here also further investigates whether the strategies used by the students in the control group had developed as a result of extra reading practice and the maturation effect.

An attempt to cluster the items by strategy using a factor analysis did not reveal a clear clustering of strategies. The clusters formed in the MAI pre-test did not carry over to the MAI post-test. No previous clustering had been attempted for this MAI questionnaire; only a subgroup of items had been clustered by Mokhtari and Reichard (2002). Therefore, the clustering was based on the researcher's own interpretation of the statements, which was informed by the literature review, and peer-reviewed by a second educational researcher. However, no formal factor analysis was carried out prior to the intervention, e.g. by analysing the data collected as part of the pilot study. This will be discussed further in Section 6.3 and Chapter 7.

The scores were calculated for each questionnaire item and then aggregated according to the strategies. Table 6.2 shows the assignment of the items to strategies. The pre-test, post-test and gain scores were calculated for each strategy and compared across the groups.

6.1.2.2.1. At baseline

The scores for each of the 16 strategies were calculated. The results of the control (N= 70) and strategy (N = 94) groups for each strategy were compared using an independent samples t-test. On average, the results showed no statistically significant difference between the groups with regard to the performance of any of the strategies (Table 6.3).

Strategy	Description	Items (Qxx)
S1	Pre-reading strategies & goal setting	1,2,23
S2	Use of prior knowledge	3,43,48
S3	Self-questioning	4,24
S4	Making inferences	20,39
S5	Visualising	21,44
S6	Rereading	5, 6, 25,45F
S7	Reading forwards/backwards	7,26,27
S8	Change of reading speed	9,28
S9	Use of context	10,11,13,29,30,31
S10	Use of external sources	14,33
S11	Summarising	15,34,35
S12	Main ideas	12,40,47F
S13	Paraphrasing	16,36
S14	Monitoring and supervising	8,32,46
S15	Evaluating	17,18,19,37,38,42
S16	Strategy awareness	22,41

Table 6.2.: Matching of items to strategies in MAI

Strategy	Group	Mean	SD	t	p-value	r
S1 - Pre-reading strategies	Control	7.59	2.10	.85	.394	.07
	Strategy	7.33	1.74			
S2 - Use of prior knowledge	Control	7.20	1.80	-.55	.583	.04
	Strategy	7.37	2.11			
S3 - Self-questioning	Control	3.73	1.65	-.81	.419	.06
	Strategy	3.94	1.61			
S4 - Making inferences	Control	4.21	1.46	-1.58	.117	.12
	Strategy	4.61	1.65			
S5 - Visualising	Control	5.70	1.54	.50	.619	.04
	Strategy	5.57	1.64			
S6 - Rereading	Control	11.49	2.30	.66	.511	.05
	Strategy	11.26	2.16			
S7 - Reading forwards/backwards	Control	8.20	1.79	1.81	.073	.14
	Strategy	7.63	2.16			
S8 - Change of reading speed	Control	5.00	1.38	.05	.962	.00
	Strategy	4.99	1.43			
S9 - Use of context	Control	14.94	3.04	1.19	.236	.09
	Strategy	14.39	2.84			
S10 - Use of external sources	Control	4.66	1.56	-.01	.992	.00
	Strategy	4.66	1.56			
S11 - Summarising	Control	6.11	1.96	-1.09	.279	.09
	Strategy	6.46	2.03			
S12 - Main ideas	Control	7.37	1.61	-.65	.519	.05
	Strategy	7.54	1.73			
S13 - Paraphrasing	Control	5.09	1.51	.68	.495	.05
	Strategy	4.91	1.63			
S14 - Monitoring and supervising	Control	8.06	1.87	.88	.379	.07
	Strategy	7.78	2.12			
S15 - Evaluating	Control	14.11	2.97	-.73	.464	.06
	Strategy	14.50	3.58			
S16 - Strategy awareness	Control	4.37	1.36	.24	.814	.02
	Strategy	4.32	1.43			

Table 6.3.: Comparison of MAI pre-test scores across groups

6.1.2.2.2. At immediate post-test

To test whether the reported use had changed after the students participated in the intervention, the group means for each strategy were compared using an independent samples t-test. The results showed a statistically significant difference in the perceived use of the five strategies (S1, S2, 24, S15, and S16) listed in Table 6.4. Note that, in all instances, the strategy group was found to report a higher perceived use of strategies.

Strategy	Group	Mean	SD	t	p-value	r
S1 - Pre-reading strategies	Control	8.29	2.09	-1.97	.051	.15
	Strategy	8.91	1.98			
S2 - Use of prior knowledge	Control	7.57	2.21	-2.02	.045	.16
	Strategy	8.22	1.91			
S4 - Making inferences	Control	4.43	1.72	-2.37	.019	.18
	Strategy	5.09	1.78			
S15 - Evaluating	Control	14.29	3.72	-2.67	.008	.20
	Strategy	15.8	3.48			
S16 - Strategy awareness	Control	4.57	1.51	-5.42	$<.001$	.39
	Strategy	5.88	1.55			

Table 6.4.: Comparison of MAI post-test scores across groups

6.1.2.2.3. Gain scores

The gain scores from both groups were compared using an independent sample t-test. The results revealed statistically significant differences between the groups regarding the use of six strategies (S1, S7, S9, S14, S15, S16) (Table 6.5).

6.2. Responses treated as categorical (binary) data

The frequencies of the responses to both versions of the questionnaire (i.e. the MAI pre-test and MAI post-test) were calculated in order to compare the students' perceptions of the use of individual strategies across the groups. Although the rating scale consisted of four levels, the responses were collapsed into binary responses to reduce the complexity and enhance the interpretability of the results. The answers

Strategy	Group	Mean	SD	t	p-value	r
S1 - Pre-reading strategies	Control	0.70	2.19	-2.49	0.014	0.19
	Strategy	1.59	2.29			
S7 - Reading forwards/backwards	Control	0.03	2.07	-3.08	0.002	0.23
	Strategy	1.14	2.43			
S9 - Use of context	Control	0.97	2.98	-2.56	0.011	0.20
	Strategy	2.23	3.23			
S14 - Monitoring and supervising	Control	0.00	2.08	-2.36	0.019	0.18
	Strategy	0.82	2.27			
S15 - Evaluating	Control	0.17	3.61	-2.02	0.045	0.16
	Strategy	1.30	3.48			
S16 - Strategy awareness	Control	0.20	1.71	-4.75	$<.001$	0.35
	Strategy	1.56	1.90			

Table 6.5.: Comparison of MAI gain scores across groups

'never or almost never' and 'only occasionally' were combined and are thought to identify students who did not report the use of reading strategies. Similarly, the answers 'usually' and 'always or almost always' were combined and are thought to identify students who reported the use of reading strategies. The frequencies of the responses on the four-point rating scale for each of the questionnaire items are reported in Appendix E.1.

6.2.1. Comparison over time

Table 6.6 shows the students' perceptions of their use of strategies by presenting the proportion of students reporting the use or non-use of reading strategies in the MAI pre-test and MAI post-test. Proportions were used instead of actual numbers of students because the size of the control and intervention groups differed (Control: N=70; Strategy: N=94). The differences in the proportions across the groups were compared using the Chi-square test, which is an appropriate test for comparing the frequency of categorical data across independent samples. The differences in the proportions within the groups over time were assessed using the McNemar test. For the Chi-square test, both test statistic and p-values are reported. For each item, the first line reports the Chi-square statistic, the second line reports the p-value. For the McNemar test, only p-values are reported.

Item	Group	% of students		McNemar test	Chi-square test	
		Pre-test	Post-test		Pre-test	Post-test
Q1	Control	39	39	1.000	3.77	5.10
	Strategy	24	56	**<.001**	**0.052**	**0.024**
Q2	Control	46	59	0.122	0.02	0.47
	Strategy	47	64	**0.017**	0.889	0.493
Q3	Control	63	66	0.845	0.18	0.10
	Strategy	60	68	0.169	0.670	0.749
Q4	Control	23	29	0.503	0.30	1.69
	Strategy	27	38	0.071	0.584	0.194
Q5	Control	86	84	1.000	0.42	0.56
	Strategy	82	88	0.286	0.516	0.456
Q6	Control	43	46	0.860	2.07	0.30
	Strategy	32	50	**0.012**	0.150	0.587
Q7	Control	69	70	0.629	3.12	5.50
	Strategy	60	84	**0.001**	0.077	**0.019**
Q8	Control	80	81	1.000	2.90	0.35
	Strategy	68	78	0.163	0.088	0.556
Q9	Control	67	69	1.000	0.73	2.17
	Strategy	61	79	**0.006**	0.392	0.141
Q10	Control	47	47	1.000	2.88	4.55
	Strategy	34	64	**<.001**	0.090	**0.033**
Q11	Control	26	30	0.701	0.70	0.40
	Strategy	20	26	0.458	0.404	0.526
Q12	Control	39	36	0.832	0.36	0.04
	Strategy	34	37	0.761	0.550	0.842
Q13	Control	64	66	1.000	8.35	1.13
	Strategy	41	73	**<.001**	**0.004**	0.287
Q14	Control	57	61	0.701	0.60	0.03
	Strategy	51	63	0.099	0.440	0.861
Q15	Control	37	43	0.572	0.18	0.82
	Strategy	40	50	0.163	0.670	0.365
Q16	Control	63	64	1.000	1.88	0.12
	Strategy	52	62	0.200	0.170	0.735
Q17	Control	36	39	0.839	0.56	1.41
	Strategy	41	48	0.441	0.453	0.235
Q18	Control	27	27	1.000	0.89	2.24
	Strategy	34	38	0.627	0.345	0.134
Q19	Control	34	37	0.845	1.80	3.63
	Strategy	45	52	0.360	0.179	0.057
Q20	Control	34	33	1.000	0.88	4.26

Continued on next page

Table 6.6 – continued from previous page

Item	Group	% of students		McNemar test	Chi-square test	
		Pre-test	Post-test		Pre-test	Post-test
	Strategy	41	49	0.337	0.348	**0.039**
Q21	Control	64	71	0.383	0.00	0.83
	Strategy	64	78	**0.029**	0.952	0.362
Q22	Control	31	34	0.864	0.12	15.06
	Strategy	34	65	**<.001**	0.725	<.001
Q23	Control	64	76	0.185	0.01	2.31
	Strategy	65	85	**0.003**	0.936	0.129
Q24	Control	24	33	0.286	0.24	0.73
	Strategy	28	39	0.080	0.627	0.392
Q25	Control	67	66	1.000	1.27	0.00
	Strategy	59	66	0.265	0.259	0.974
Q26	Control	24	30	0.541	4.13	0.93
	Strategy	39	37	0.868	**0.042**	0.334
Q27	Control	80	70	0.167	9.24	5.45
	Strategy	57	85	**<.001**	**0.002**	**0.020**
Q28	Control	36	31	0.690	0.05	0.40
	Strategy	34	36	0.864	0.824	0.526
Q29	Control	49	69	**0.007**	0.01	0.05
	Strategy	48	70	**0.001**	0.929	0.821
Q30	Control	61	70	0.327	0.55	0.03
	Strategy	67	71	0.618	0.459	0.859
Q31	Control	56	57	1.000	0.60	5.45
	Strategy	62	74	**0.050**	0.440	**0.020**
Q32	Control	47	43	0.728	0.25	6.41
	Strategy	51	63	0.126	0.619	**0.011**
Q33	Control	33	39	0.557	0.00	0.03
	Strategy	33	37	0.618	0.987	0.861
Q34	Control	16	24	0.180	3.28	0.11
	Strategy	28	27	1.000	0.070	0.737
Q35	Control	33	41	0.263	0.09	0.17
	Strategy	35	45	0.188	0.764	0.678
Q36	Control	37	56	**0.019**	0.08	0.05
	Strategy	39	57	**0.014**	0.773	0.825
Q37	Control	36	46	0.210	1.34	3.10
	Strategy	45	60	**0.024**	0.248	0.078
Q38	Control	57	56	1.000	0.14	1.42
	Strategy	54	65	0.076	0.713	0.233
Q39	Control	31	41	0.143	2.11	3.10
	Strategy	43	55	0.073	0.146	0.078

Continued on next page

Table 6.6 – continued from previous page

		% of students		McNemar test	Chi-square test	
Item	Group	Pre-test	Post-test		Pre-test	Post-test
Q40	Control	43	49	0.557	0.25	5.00
	Strategy	47	66	**0.011**	0.615	**0.025**
Q41	Control	39	44	0.503	1.76	20.66
	Strategy	29	79	**<.001**	0.184	**<.001**
Q42	Control	64	61	0.845	0.00	3.77
	Strategy	64	76	0.090	0.952	**0.052**
Q43	Control	27	33	0.557	0.05	4.26
	Strategy	29	49	**0.002**	0.824	**0.039**
Q44	Control	67	64	0.824	0.73	0.13
	Strategy	61	67	0.377	0.392	0.715
Q45F	Control	71	66	0.481	0.08	0.06
	Strategy	73	64	0.200	0.779	0.803
Q46	Control	50	44	0.585	0.90	3.77
	Strategy	43	60	**0.007**	0.344	**0.052**
Q47F	Control	76	59	**0.012**	0.61	0.17
	Strategy	70	55	**0.034**	0.435	0.678
Q48	Control	56	56	1.000	0.05	2.18
	Strategy	57	67	0.163	0.825	0.140

Table 6.6.: Comparison of MAI responses at pre-test and post-test

Table 6.6 should be interpreted as follows: for example, in question Q1, the proportion of students in the control group agreeing with the statement was the same for the MAI pre-test and the MAI post-test (39%) (p=1.000). Students in the strategy group, however, were more likely to report the use of strategies in the MAI post-test (56%) than in the MAI pre-test (24%), and this difference was statistically significant (p<.001). The difference across groups regarding the proportion of students in agreement was borderline statistically significant in the MAI pre-test (39% vs. 24%, p=.052), with the strategy group being lower than the control group, and in the MAI post-test (39% vs. 56%, p=.024), with the strategy group being higher than the control group. The p-values highlighted in bold in Table 6.6 are those that were statistically significant at $\alpha < .05$ or borderline significant ($0.05 \leq \alpha < 0.055$).

Figure 6.1 illustrates these patterns visually. For each question, the proportions of students reporting strategy use before and after the intervention were plotted. The differences in levels reflect the change in usage patterns over time.

The results from the McNemar tests showed statistically significant changes in agreement with regard to 17 items for the students in the strategy group (described

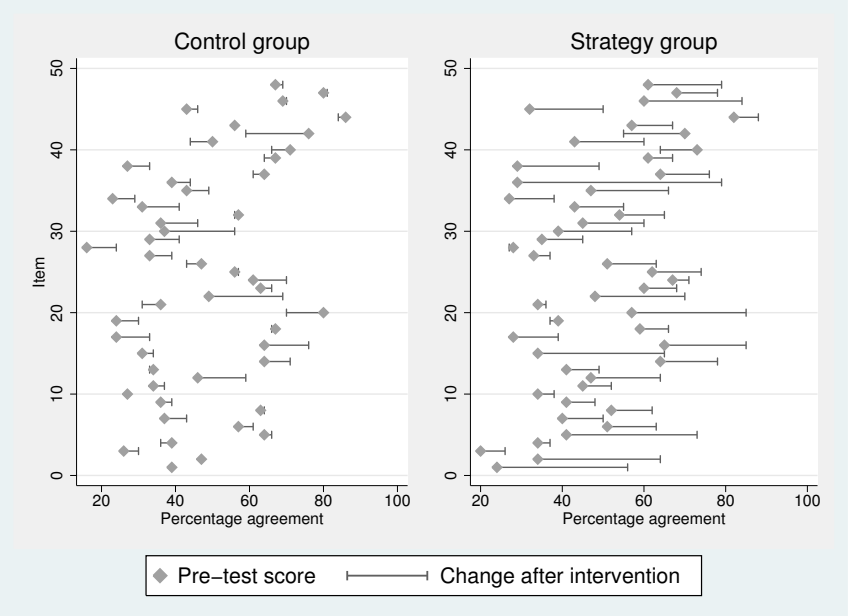


Figure 6.1.: Changes in reported awareness and use of strategies

in full in Table 6.7): Q1, Q2, Q6, Q7, Q9, Q10, Q13, Q21, Q22, Q23, Q27, Q31, Q37, Q40, Q41, Q43, and Q46. It is worth noting that, in the control group, the scores did not increase statistically significantly for any of these items. In contrast, for items Q29, Q36, and Q47F, the change between the MAI pre-test and MAI post-test was statistically significant in both groups (Table 6.8).

Item	Item description	Strategy
Q1	I think about the reason for reading the text before reading it to help me during my reading.	S1 - Pre-reading strategies
Q2	I quickly preview the text to become aware of its organization, structure, length and other characteristics	S1 - Pre-reading strategies
Q6	When I read a text, I read it several times to improve my comprehension	S6 - Rereading
Q7	When I find a phrase that I do not understand, I read before and after the phrase to help me to understand the meaning of the phrase	S6 - Rereading
Q9	I adjust my reading speed according to the text I'm reading (if it is a difficult text, I read it slowly) to try to comprehend it	S8 - Change of reading speed
Q10	I use the clues (i.e. keywords) in the text to help me better to understand what I'm reading	S9 - Use of context
Q13	I use typographical aids (i.e. boldface and italics) to help me to identify important information	S9 - Use of context
Q21	I try to visualise information to help me to remember what I'm reading.	S5 - Visualise
Q22	When I read, I have in mind the reading strategies that can help me to understand the text	S16 - Strategy awareness
Q23	Before reading the text, I skim the text, title, and pictures to see what is about	S1 - Pre-reading strategies
Q27	If there's a word that I do not understand, I read the words before and after it to help me to understand it	S6 - Rereading
Q31	When I find a paragraph difficult to understand, I keep reading to see whether it becomes clear later on	S9 - Use of context
Q37	I discuss my reading with others to check my understanding	S15 - Evaluate
Q40	When I read, I look for the most important part of the text	S1 - Pre-reading strategies
Q41	I know when to use the reading strategies	S16 - Strategy awareness
Q43	I try to link the information in the text with my other subjects at school	S2 - Use of prior knowledge
Q46	I check my understanding of what I've read and ask myself which concepts I haven't understood	S14 - Monitor and supervise

Table 6.7.: MAI items for which the strategy group showed statistically significantly increased awareness

Item	Item description	Strategy
Q29	I use the tables, figures, and pictures in the text to increase my understanding	S9 - Use of context
Q36	I paraphrase (restate ideas in my own words) in order better to understand what I'm reading	S13 - Paraphrasing
Q47F	When I read, I highlight the details of the text (reversed item)	S12 - Main Ideas

Table 6.8.: MAI items for which both groups showed statistically significantly increased awareness

After the intervention, more students in both groups reported that they were in agreement with item Q29, which referred to the use of charts, graphs and pictures to improve comprehension (n=13, a change of +18% in the control group; n=21, a change of +23% in the strategy group). The increase in the control group might be explained by the natural development of strategies that the students experienced due to the practice effect (i.e. increase in reading time). Although the students in the control group worked with the same texts as the students in the strategy group, both groups experienced an increment in the time spent on reading activities at school and the type of text to which they were normally exposed. Also, the increased exposure to texts that intentionally use tables and questions that exploit this strategy (i.e. use of context) might have indirectly helped the students to develop the strategy by providing opportunities for practice.

Item Q36 also shows increases in the proportion of students reporting the strategy in both groups (n=12, a change of +12% in the control group; n=17, a change of +18% in the strategy group). This item referred to the paraphrasing of information. This increment in the control group might be explained by the fact that the activities developed with this group included discussion of the text. The students might have found it easier to explain their ideas using their own words; thus the change in the reports of using this strategy.

Item 47F was a reversed item. The students had to report whether they highlighted details in a text. The proportion of students reporting agreement with the associated statement dropped in both groups (n=11, a change of -16% in the control group; n=15, a change of -16% in the strategy group). These results show that the students were emphasising the main ideas in the texts. For the students in the control group, this might be a side effect of practice during the discussions.

Despite the increase in the number of students in both groups agreeing with the statements made in Q29, Q36 and Q47F after the intervention, it is worth noting that

the control group never improved more than the strategy group in any of these items.

6.2.2. Comparison between groups at baseline

The previous analyses showed statistically significant increases after the intervention in the reported agreement with the statements made regarding 20 items. However, these results do not take into account any pre-existing differences between the groups. To test whether these increases are the result of the intervention rather than pre-existing differences, it is necessary to compare the performance of both groups at baseline. The results from a Chi-square test carried out on the MAI pre-test showed statistically significant differences between the groups regarding four items (Q1, Q13, Q26, and Q27), with students in the control group generally showing more agreement than their peers in the strategy group (Table 6.9). Moreover, the students' perceptions at baseline were generally similar and not statistically significantly different. Baseline differences cannot therefore explain the larger gains in reported use observed in the strategy group compared to the control group.

Item	Item description	Strategy
Q1	I think about the reason for reading the text before reading it to help me during my reading.	S1 - Pre-reading strategies
Q13	I use typographical aids (i.e. boldface and italics) to help me to identify important information.	S9 - Use of context
Q26	I go back and forth in the text to find the relationships among the ideas in it.	S7 - Reading forwards / backwards
Q27	If theres a word that I do not understand, I read the words before and after it to help me to understand.	S6 - Rereading

Table 6.9.: MAI items with statistically significantly group differences at baseline

The results from the Chi-square performed on the MAI post-test yielded statistically significant differences between the groups regarding 12 items (Q1, Q7, Q10, Q22, Q27, Q31, Q32, Q40, Q41, Q42, Q43, Q46), whereby the students in the strategy group were more likely to report agreement. Therefore, it could be said that there is an association between group membership and strategy-use.

6.2.3. Confounding variables

Similar to Chapter 5 regression analysis was used to test whether a range of potentially confounding variables (number of books that students had at home, the amount

of time spent reading for pleasure, socio-economic status (SES), and attendance rate) were associated with the aggregate MAI post-test scores. This served to answer the following questions:

1. Are the four variables (number of books at home, reading time, SES, and attendance rate) statistically significantly associated with variations in the MAI post-test after accounting for the group assignment and MAI pre-test scores?

2. If so, how much of the variability in the scores was accounted for by these variables?

3. Does the observed association between treatment assignment and MAI post-test scores, conditional on MAI pre-test scores, change once these additional variables are controlled for?

Two linear regression models were estimated using the ordinary least squares estimator (Table 6.10). The first model included only the baseline (i.e. MAI pre-test) scores and an indicator variable for the group membership (i.e. control and strategy). The second model included the same variables as the previous model plus the four variables for the number of books at home, reading time, SES, and attendance rate. The dependent variable in both models was the MAI post-test score.

The results of the regression indicated that MAI pre-test scores and group membership accounted for 38% of the variance in the MAI post-test scores (model 1) (R^2=.377, F(2,163)=48.64, p<.001) and the other variables (model 2) accounted for an additional 6.5% of the variance (R^2=.065, F(11,163)=10.95, p<.001).

	Immediate post-test			Delayed post-test		
Variable	β	SE	p-value	β	SE	p-value
Constant	40.59	8.96	<.001	33.04	13.62	0.016
MAI Pre-test	0.69	0.07	<.001	0.62	0.07	<.001
*Control group**	-	-	-	-	-	-
Strategy group	9.07	2.57	0.001	9.60	2.59	<.001
Socio-economic status				0.30	0.36	0.405
Attendance				0.07	0.09	0.477
*Reading 0 min**				-	-	-
Reading i30'				9.84	3.23	0.003
Reading 30 - 60'				11.98	3.95	0.003
Reading 1 - 2 hours				12.94	4.94	0.010
Reading ¿ 2 hours				16.72	6.40	0.010
*Books $\leq$10**				-	-	-
Books 11 - 25				-5.24	3.60	0.148
Books 26 - 100				-7.74	3.73	0.040
Books ¿100				-5.49	4.66	0.241

(*) Base category; β = beta; SE = standard error

Table 6.10.: Results of linear regression with MAI pre-test scores and further potential confouding variables as covariates

The estimated coefficient β indicates the effect of the respective variable on the outcome when all other variables are held constant. The results from the first model showed that the MAI pre-test scores were statistically significantly associated with the MAI post-test scores. For every unit by which the MAI pre-test score increases, the MAI post-test score is expected to increase by approximately .69 units. The group membership variable was statistically significantly associated with the MAI post-test scores. Being in the strategy group was associated with 9.07 units higher MAI post-test scores.

In the second model, variation in the MAI post-test scores was again statistically significantly associated with the MAI pre-test scores and the group membership, and additionally with the time for which the students read for pleasure. The results suggested that students who read more than two hours per day for pleasure will achieve 16.72 units higher scores in the MAI post-test than those who did not read for pleasure at all (the base category). Note that the coefficient on the MAI pre-test score is very similar in both models as well as the group effect. This suggests that the other factors are not confounding the estimated association between the MAI post-test scores and pre-test scores or group membership, but are merely additional

explanatory factors.

The results from the multiple regression indicate that the extra time that the students read for pleasure, group membership and, to some extent, the number of books affect the students' reported strategy use.

6.2.4. Lessons from MAI

Responses of the self-reported questionnaire were analysed treating data as interval and categorical measures. Results from analyses treating responses as interval data showed that both groups were similar at baseline and that after the intervention students in the strategy group reported a higher strategy use. The analysis of gain scores showed that students in the strategy group reported statistically significantly higher increases in usage than students in the control group. An attempt to analyse the items by strategy did not reveal a clear clustering of strategies (see Section 6.3). Other variables, such as number of books at home and time spent reading for pleasure, are statistically significantly associated with the MAI post-test scores, but they do not confound the relationship between group membership and gains in MAI scores.

The analysis in which data were treated as categorical identified differences between groups in four items before the intervention; with a higher proportion of students in the control group reporting strategy use. After the intervention, statistically significant differences were found in 12 items, in all of which the strategy group reported higher use. Results of comparisons over time showed that students in the strategy group had increased their strategy use in 20 items while students in the control group only reported increased usage in 3 items.

Both types of analysis – treating data as interval or categorical – yielded similar results. The students in the strategy group reported a higher use of strategies after the intervention than their peers in the control group. The baseline differences were small and generally statistically insignificant. Hence, it might be said that the students in the strategy group became (more) aware of the strategies and began to use them more as a result of their participation in the intervention.

6.3. Factor analysis

The results of the previous sections showed that the teaching of reading strategies during the intervention had a positive effect on the students' self-reported use of strategies. However, these findings are based on the researcher's own clustering of

items. An exploratory factor analysis (EFA) was carried out to identify how the items cluster and see whether this clustering was similar before and after the intervention (A. Field, 2009). Mokhtari and Reichard (2002), for instance, found that MARSI revealed three different types of strategies (i.e. factors), which they labelled global, problem-solving, and support strategies. For the metacognitive questionnaire used in the current study, a scree plot and an analysis of the eigenvalues[1] revealed that 15 factors should be retained. The EFA was carried out using all MAI data collected, i.e. the MAI pre-test and MAI post-test for both groups. The results revealed that the strategies did not cluster as expected and no cluster pattern could be identified when retaining 15 factors (Appendix E.2). A further attempt was made to replicate the factor analysis carried out by Mokhtari and Reichard (2002) using only data on the 30 items included in MARSI. In contrast to previous results reported by Mokhtari and Reichard (2002) for MARSI, no clear cluster pattern of strategies was identified in the current data. This suggests that the different items may measure different factors and that an attempt to cluster the data into strategies was likely to fail.

6.4. Verbal reports

The analyses of the MAI data showed an increased use of strategies amongst the students in the strategy group. These results, however, might be biased because they are based on the students' subjective self-assessment rather than an objective measure of strategy use. Therefore, the verbal protocols ('think alouds') were analysed to provide a more objective assessment of reading strategy usage. The verbal protocols were included in this study to obtain data on the processes followed by the students while reading unfamiliar texts. The subsamples of students from each group were asked to perform reading tasks before and after the intervention. The students' utterances while performing the reading tasks were transcribed and coded for further analysis.

The following section describes the different analyses carried out with the data gathered as part of the verbal protocols. First, the students' think alouds were audio recorded and later transcribed. The analyses of the transcripts served to establish whether the students reported using the strategies to solve reading tasks when faced with unfamiliar texts. Coding criteria, which reflected the strategies taught in the intervention, were developed to categorise the students' responses for further analysis.

[1] Following the suggestion of Kaiser (1960), eigenvalues above 1 were considered substantial and the corresponding factors were retained in the EFA.

Then, the students' ability to detect incongruences was assessed. The number of incongruences detected was used to monitor the students' use of metacognitive strategies. Comparisons between the groups served to answer the third research question (Q3):

- Q3: Does the intervention increase the reported 'online' use of reading comprehension strategies when reading unfamiliar texts?

6.4.1. Coding

6.4.1.1. First coding

A codebook was developed based on the preliminary analysis of the transcripts (i.e. inductive coding) and the strategies taught during the intervention. It contained an explanation for each strategy and examples of the coding. Nvivo software version 19 was used to carry out the coding task. The interviews were coded following two main criteria:

1. *explicit code* - the strategy was explicitly named by the student
2. *code in use* - the context showed a clear example of the strategies being used by the student

The first codebook covered ten strategies[2] derived from the intervention (see Appendix C.13) and four categories, which arose from the transcript and the initial coding:

1. *lack of comprehension* - indicated segments where the students expressed their lack of comprehension of the text or when the discussions and/or students' questions clearly showed a misunderstanding of the text
2. question misunderstanding - indicated extracts where students were unable to answer questions set in the reading task because they did not understand what was required of them
3. *text reference* - indicated that the students were not consulting the text to carry out the tasks; instead, they used their previous knowledge or memorised

[2]Some strategies were grouped together because it was difficult to differentiate between them from the verbal reports (e.g. the three strategies S7 - rereading, S8 - readining forwards/backwards and S9 - change of reading speed)

segments of the text. This code included utterances where the students explicitly asked whether they were allowed to check the text again

4. *vocabulary* - indicated parts of the interviews where the students encountered comprehension problems due to poor vocabulary, i.e. the students did not know the meaning of an important word

A copy of the transcripts and the codebook was shared with a second coder[3]. The details about the groups and the type of test (i.e. pre-test or post-test) were masked. The intercoder agreement was calculated using Cohen's Kappa. The assessment showed a 'fair" agreement (κ=.33) (Bernard and Ryan, 2010, p. 30). Discrepancies were discussed and the codebook was amended to clarify any unclear points. Coding criteria and changes to the codebook were discussed with a third educational researcher.

6.4.1.2. Second coding

The improved version of the codebook was again shared with the second coder to assure that he agreed with the changes made. A further 10% of the sample was recoded using the new version of the codebook. The intercoder agreement was κ=.66 which is considered substantial (Bernard and Ryan, 2010, p. 30). Table 6.11 shows examples of the data coding for some of the strategies.

[3]The second coder was a Mexican researcher with experience of qualitative analysis.

Strategy	Name given in the intervention	Examples (Translated from Spanish)
Pre-reading activities	Why do I have to read? (¿Por qué leer?)	R: What are you doing? S105: First, I check what I have to do and then look at it in the text. S146: Well, I always skip the text first and read the questions and then look for the answer in the text.
Scanning	Quick review (Reviso rápido)	I quickly check the paragraphs to see which one is more [...]
	Locate	R: What did you do to find the answer? S101: Look for it in this part [...]
	What is it about? (¿De qué se trata?)	R: What do you mean by analyse? S: Well, I analyse what the text is about
Visualise, summarise, make inferences	Review and Imagine (Repaso e imagino)	S146: I think it must have teeth... It must be as any other bird with two wings and tornasol colour
Rereading	Rereading (Vuelvo a leer)	S171: I read again, but I checked, and that is not the correct answer
	Read backwards and forwards (Leo atrás y adelante)	S106: Go back and forwards to check if I can understand

R = Researcher, S = Student ID

Table 6.11.: Verbal protocol coding examples

6.4.2. Transcript analysis

Coded extracts of the verbal protocols were analysed using content analysis. This analysis served to see whether the intervention had increased the students' use of reading strategies when reading unfamiliar texts. Each time a student used a strategy, named it or described it in their own words, it was counted as a unit or 'token'. The number of tokens was counted for the reading tasks before and after the intervention for both groups. This information was used to create a case-by-variable matrix, which showed how many times participants from both groups used each strategy during the reading tasks (Appendix E.3). Descriptive analysis of the matrix showed that, before the intervention, the students in the strategy group registered two more tokens (N=11) than those in the control group (N=9)(Table 6.12). The number of times the students in the strategy group registered tokens after the intervention increased

substantially more (to N=143) than in the control group (to N=10).

	Pre-test	Post-test
Control	9	10
Strategy	11	143

Table 6.12.: Frequency of registered tokens in verbal report

These results reflected the frequency of strategy-use but did not show the proportion of students using the strategies. Before the intervention, six of nineteen students (31.6%) in the control group used strategies whilst, in the strategy group, nine of fourty students (22.5%) used strategies (Table 6.13). After the intervention, a larger proportion of students in the strategy group used strategies (29 out of 40 - 72.5%) than in the control group (8 out of 19 - 42.1%).

	Pre-test	Post-test
Control	31.6%	42.1%
Strategy	22.5%	72.5%

Table 6.13.: Percentage of students registering tokens in verbal report

The tokens were also aggregated by strategy in order to rank strategies by usage (Table 6.14). The results showed that the students in the control group used three strategies before the intervention: scanning to locate, rereading to understand, and summarisation. None of the other strategies were used at all (zero tokens). After the intervention, these students were rereading to understand about six times more often than before the intervention. For the remaining two strategies (i.e. scanning to locate and summarising), no major change was observed.

The students in the strategy group also used only three strategies before the intervention, two of which were used by the control group (scanning to locate, rereading to understand) and one that was not (reviewing and imagining). After the intervention, the students were using 10 strategies. Five strategies were used more than 15 times: quickly scanning, scanning to locate, reviewing and imagining, rereading to understand, and reading forwards and backwards.

The previous results showed that students in the strategy group were actively using the strategies learnt during the intervention. To verify the use of metacognitive strategies further, specifically monitoring, the students were asked to identify

Strategy	Measurement	Group	
		Control (N=19)	Strategy (N=40)
Why I have to read?	Pre-test		
	Post-test		5
Quick review	Pre-test		
	Post-test		17
Locate Scan	Pre-test	7	9
	Post-test	3	27
What is it about?	Pre-test		
	Post-test		1
Review and Imagine	Pre-test		1
	Post-test	1	31
Rereading	Pre-test	1	1
	Post-test	6	19
Read backwards and forwards	Pre-test		
	Post-test		20
Speed change	Pre-test		
	Post-test		4
Use of context	Pre-test		
	Post-test		9
Summarise	Pre-test	1	
	Post-test		10

Table 6.14.: Number of students registering tokens on individual strategies before and after the intervention

incongruences in the text.

6.4.3. Discrepancies detection

Metacognition is the process that helps students to gain an awareness of their cognitive processes (see Chapter 3). In the field of reading, it has been described as the process that helps students to know when and why to use reading strategies. Metacognitive processes have been linked with the monitoring, supervision and evaluation of comprehension (Garner, 1987; Hacker, 1998). To measure monitoring and evaluation, some studies inserted incongruences into texts and asked the students to read the texts. Zinar (2000) suggested that, if students are able to identify such incongruences, it means that they are monitoring their comprehension and supervising their understanding of the text. The insertion of incongruences was adopted in the current study as one indication of whether the participants were genuinely paying attention to the texts (under the assumption that participants who were disengaged would be less likely to spot anomalies).

Each text in the verbal protocols contained five incongruences. The number of incongruences identified by the students in each group was not normally distributed (see Table 6.15). Therefore, changes in the number of incongruences detected over time were assessed using the non-parametric Wilcoxon signed-rank test. The number of incongruences at pre-test and post-test were compared across the groups using the Mann-Whitney test.

	Pre-test (Text 1)		Post-test (Text 1)		Post-test (Text 2)	
# Errors	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Control group						
0	9	52.9	10	58.8	5	29.4
1	5	29.4	1	5.9	4	23.6
2	0	0	3	17.6	4	23.6
3	3	17.6	2	11.8	2	11.8
4	0	0	1	5.9	2	11.8
5	0	0	0	0	0	0
Strategy group						
0	20	66.7	17	56.7	13	43.3
1	7	23.3	8	26.7	5	16.7
2	3	10	2	6.7	4	13.3
3	0	0	1	3.3	7	23.3
4	0	0	1	3.3	1	3.3
5	0	0	1	3.3	0	0

Table 6.15.: Number of incongruences detected

The results showed a statistically significant improvement (z=-2.05, p=.041, r=.62) in the number of incongruences detected in the post-test (mean = 1.21, SD = 2.02) compared to the pre-test (mean = 0.37, SD = 0.68) by students in the control group. Similarly, the students in the strategy group detected more incongruences in the post-test (mean=5.51, SD=5.13) than in the pre-test (mean=.40, SD=.66) and this difference was also statistically significant (z=-4.84, p<.001, r=.87).

The results from the Mann-Whitney test showed no statistically significant difference between the control group (mean=.37, SD=.68) and the strategy group (mean=.40, SD=.66) at baseline (U=395.5, p=.803, r=.03). Comparisons after the intervention showed a statistically significant difference between the control group (mean=1.21, SD=2.02) and the strategy group (mean=5.51, SD=5.13), with the strategy group identifying more incongruences (U=210, p=.002, r=.39). The gain scores were calculated and compared between the groups. On average, the strategy group achieved larger gains (mean=5.12, SD=2.00) than the control group (mean=.84, SD=5.12). This difference was statistically significant (U=192, p=.001, r=.43).

6.4.4. Lessons from the verbal protocols

The results from the analyses of the verbal protocols showed that students in both groups used strategies before taking part in the intervention. This finding was expected because research has shown that even poor readers use some reading strategies (although not necessarily efficiently) (Garner and Kraus, 1981). Scanning was the most commonly used strategy. This strategy is practised in school contexts and students are encouraged to use it in the question-answer exercises provided to assess reading comprehension. However, the analysis of the number of strategies and the number of times these strategies were used before the intervention suggest a poor prior use of strategies. The results after the intervention showed an increase in the number of times that the students use strategies, and an increase in the range of strategies used. The increases were statistically significant in the strategy group but not in the control group.

6.5. Analyses of combined instruments

The main aim of the second part of the chapter is to combine the results of the instruments in order to answer research question Q4:

- Q4: Is there a relationship between the reported use of strategies and performance on the reading comprehension test?

Spearman's correlation coefficient (ρ) was calculated for four different gain scores in order to test whether there was a relationship between reading comprehension scores and the students' perceptions of strategy-use. Two gain scores were calculated for the reading comprehension test (i.e. G1 = immediate post-test - pre-test; G2 = delayed post-test - pre-test) and one gain score each was calculated for the MAI (i.e. MAI gain = MAI post-test - MAI pre-test) and the number of tokens of reported strategy use in the verbal report (i.e. VR gain = VR post-test VR pre-test). The results from the correlation between G1 and MAI gain did not show a statistically significant relationship for the control group (ρ=-.096, p=.431) or the strategy group (ρ=-.104, p=.325). Similar results were found for the correlation between G2 and MAI gain for the control group (ρ=-.070, p=.590) and the strategy group (ρ=.042, p=.711). These results should be interpreted with care because the analysis included self-reported data and so the students' responses might not reflect what they actually do. Therefore, gains from the reading comprehension test were also compared with gains in the verbal reports. The correlation between G1 and the VR gain scores were statistically insignificant for the control group (ρ=.198, p=.430) and the strategy group (ρ=.176, p=.284). Similar results were found for the correlation between G2 and verbal report gain scores: ρ=-.100, p=.746 for the control group, and ρ=.150, p=.389 for the strategy group. In summary, there were no statistically significant correlations between any of the gain scores derived from the different instruments.

7. Discussion

The aim of this study was to test the effectiveness of the direct teaching of reading comprehension strategies. The study makes several original contributions to the literature: first, it investigates the effectiveness of the direct teaching of reading comprehension strategies on self-reported strategy use, observed strategy use, and actual comprehension simultaneously, and corroborates the information gathered. Second, the results for the intervention group were compared with those of the control group, who experienced the same increase in reading activity using the same materials, thereby ruling out, as far as possible, the possibility that changes in the reading activities may have caused observed improvements in comprehension and strategy use. Finally, the context of this study differed from that of studies which have previously tested strategy instruction i.e. in a relatively low literacy context, with Spanish speaking adolescents (see Chapter 3). Five instruments were designed and data were collected using these instruments in order to answer the five research questions. This chapter brings together the results of Chapters 5 and 6 and provides a deeper understanding of the findings. The discussion is organised according to each instrument.

7.1. Discussion of Word Level Study

The WLS was divided into three sections. The results from the word identification test indicated that the students were able to identify over 90% of the words, which is unlikely to lead to comprehension problems due to decoding difficulties (Paris and Hamilton, 2008). This percentage was the same regardless of the number of syllables or the frequency of the word, indicating that this finding is robust for different test specifications. There were no statistically significant differences between the strategy and control groups.

The word knowledge test (comprising the vocabulary recognition and similarities tests) was analysed in two ways: using overall scores and normed rating scales from the WISC-R test. For the similarities test, the results from both analyses yielded

identical results, suggesting that the students knew the meaning of the words. The students in the strategy group performed slightly better than those in the control group but this difference was not statistically significant. The results of the vocabulary test, however, differed depending on the method of assessment. When the raw data were used to make comparisons (i.e. based on an overall score, which was obtained by equal weighting, i.e. adding up points), the results indicated statistically significant better performance in the strategy group compared to the control group. This difference, however, was not significant when the comparisons were based on the WISC-R rating. As before, the overall performance in the vocabulary test suggested that the students had a sufficient knowledge of the word meanings to participate successfully in the intervention.

Three points warrant further discussion: first, the choice of assessment type - overall score vs. WISC-R rating - affected the interpretation of the results. While one test reported a statistical difference between the groups, the other reported homogeneous performance by both groups. This difference might be explained by the way in which the scores are interpreted. WISC-R scales have been validated for comparisons with the general population of students of the same age. These validated rating tables group two or three different point scores into one level. In contrast, the overall scores are continuous and allow the measurement of smaller, possibly non-significant, differences than is possible when using the grouped ranking approach. It is possible that the grouping of scores into a WISC-R rating reduced the variability of the data and, as a result, led to non-significant test results.

Second, however, the finding that the students in the strategy group possessed a better raw vocabulary recognition score should be taken into consideration when interpreting the results. Research has shown that a high level of vocabulary knowledge might lead to better comprehension (Johnston, 1984; Beck, McKeown and Kucan, 2004). Hence, the relative advantage of the strategy group over the control group might at least partially explain the higher gains that the strategy group realised after the intervention. Put differently, it is possible that differences in vocabulary knowledge at baseline contributed to the observed changes in reading comprehension rather than the intervention itself. However, in this case, one would expect to observe a higher immediate gain in the strategy group than the control group, which was not the case. This hypothesis might have been studied further by controlling for the effect of prior vocabulary knowledge as a confounder in the analysis of covariance reported in Chapter 5. However, the number of students in the subsample (Strategy: N=43; Control: N=19) was deemed too small to provide valid results from a statistical

analysis. Coupled with the fact that the results based on the WISC-R scales showed no statistically significant difference between groups, no such analysis was carried out.

Third, the students in both groups performed better in the similarities test compared to the vocabulary recognition test. This difference may be inherent to the tests and reflect differences in their degrees of difficulty. Alternatively, it might be a result of the novelty of the activity. The students might be unaccustomed to defining words in isolation, as required by the recognition test. Also, the students' oral verbal ability might have prevented them from defining the words appropriately. It seemed to be slightly easier for the students to understand the meaning of words that were placed within a context (i.e. in the similarities test). This indicates that the students might have made use of clues from the context in order to infer the correct meaning of the words. Hence, the students would be unconsciously using reading strategies before the intervention. However, because this study was designed to elicit *changes* in reading comprehension - and so, indirectly, strategy use - due to the intervention, rather than the total levels of strategy use, any strategy use before the intervention is unlikely to bias the results as long as this is comparable across the groups. Given that the students in both groups performed similarly on the similarities test, there is no reason to assume that prior strategy use, especially with respect to the use of context' strategy, differed between the groups. Unfortunately, no data were collected to investigate further the differences in performance between the similarities and word recognition tests, and this issue was deemed as lying outside the scope of the study.

7.2. Discussion of the reading comprehension test

The results of the analysis of variance using the three reading comprehension test scores (pre-test, immediate post-test and delayed post-test) as outcomes indicated that the effect of group membership itself was not significant, i.e. the level of comprehension did not differ between groups when considering all time points jointly and the group scores averaged across time. However, the results yielded significant changes in the students' comprehension over time i.e. when the gain scores were analysed. Both groups initially improved their comprehension (i.e. at the immediate post-test) and the students in the strategy group maintained their gains over time (i.e. at delayed post-test). These changes were significant for the students in the strategy group, as confirmed by the ANCOVA. No substantial effect of other

student characteristics, such as socio-economic status, reading practice at home or attendance rate, on comprehension test scores was found.

Positive results from short interventions have been criticised as being unreliable. Pressley (2000) argued that such results might be due to the novelty of the intervention or to the learning effect when improvement is assessed with two or more identical comprehension tests. He suggested delivering longer interventions (6 months to a year) to avoid this problem and to reassess the students' knowledge to look for maintenance of the results. By delivering longer interventions, the students will have more opportunities to practise the strategies learnt. This period of practice is also an important component in information-processing models. For example, Mitchell and Myles (2004) argue that learning a skill means moving from a conscious process to an automatic, unconscious one. To achieve automatisation, the action needs to be practised *"through repeated activation, [so that] sequences first produced by controlled processing become automatic"* (Mitchell and Myles, 2004, p. 101). Because short interventions lack this period of practice, automatisation is unlikely to occur. The present study followed Pressley's argumentation (Pressley, 2000), employed a long intervention (4 months follow-up) and reassessed both groups twice after the end of the intervention using different, validated questionnaire versions. The students were provided with a period of time in which to acquire the strategies and to transfer their new knowledge to new contexts. Yet, despite the initial improvement in both groups, only the strategy group maintained the gains, indicating that the intervention was effective in improving and maintaining comprehension over time. However, whether the follow-up period was sufficiently long for all of the benefits of the intervention to materialise cannot be determined with any certainty. The follow-up time was determined by the school authorities and was outside the control of the researcher.

The students' comprehension, measured by the gain scores, improved among not only those in the strategy group but also those in the control group, at least when measured immediately after the intervention. This improvement might be explained by several factors: first, when the intervention was designed, it also included activities targeted at the control group. This reflects the pragmatic trial nature of this study. Moore, A. Graham and Diamond (2003) state that a *"control treatment would in itself be a much enhanced version of the treatment normally available to participants in the absence of a trial"* (Moore, A. Graham and Diamond, 2003, p. 681). In this study, the students in the control group served as an active control group. They were actively engaged in performing reading activities during the intervention, discussing the texts and working in groups to answer the same questions as the strategy group. This

type of activity might have affected the students' performance in the control group. Research has shown that discussions promote deeper engagement with the ideas presented in the text (Tierney, Readence and Dishner, 2005; Beck, McKeown and Kucan, 2004). The interactions and experiences of the students in the control group served as extra practice, which might have resulted in comprehension gains for the control group. Therefore, even without any strategy instruction, some of the students might benefit from this interaction and extra practice, i.e. being part of the study. In sum, the discussions about the topics in order to answer the questions might also have benefited the control group so that the observed difference between the strategy and control group was smaller than would have been the case if there had been an absence of any increase in interaction and practice.

Second, another reason for the comprehension improvement in the control group is the extra exposure to reading. During the intervention, the students in both the strategy and control groups experienced an increase in reading time. Pressley (2000) suggests that one way to improve comprehension ability is to encourage extensive reading. He states that the continuous exposure to texts might result in increased exposure to words. This word exposure then leads to the acquisition of new vocabulary, which indirectly improves comprehension. It is worth noting that the basis of comprehension is the fluent, automatic recognition of words (LaBerge and Samuels, 1974). Therefore, the students in the control group might have improved because they were exposed to more extensive reading and new vocabulary during the intervention.

Finally, it could be argued that both groups might have improved at the immediate post-test because of general maturation, extraneous variables outside the experimental comparison, or the test effect (e.g. Marsden and C. Torgerson, 2012). Regardless of any intervention, students will generally show improvement in their skills, due to the maturation effect. Over the period of an intervention, or even without it, students are faced with different experiences and extra practice. All of this knowledge helps them to mature sufficiently to acquire their own knowledge without any further explanations. Although randomisation can help to control the students' individual variables, there are extraneous variables, such as the extra practice that students engage in outside school, which might influence their improvement. All of the stimulus and input outside the classroom then constitute factors which might affect the students' comprehension. The test-retest effect might also explain the gains at the immediate post-test. This effect is observed generally when a test is repeated within a short period of time. It is argued that the students might remember the

questions or have become familiar with this type of test. In this study however, three versions were designed in order to avoid this effect. Also, the difficulty of the tests was assessed during the pilot to ensure the same level of difficulty across the texts.

This study featured an active control group. The advantage of this approach is that, by exposing both groups to the same amount of practice, the observed effect in the strategy group is more likely to reflect the real additional benefit of the direct teaching of reading comprehension strategies. However, the study lacked a non-active control (i.e. a test-only group). Hence, none of the possible explanations presented above can be tested empirically using the available data. Yet, it is worth noting that these explanations fail to explain why the strategy group maintained their gains until the delayed post-test while the control group did not.

An ANCOVA was carried out to measure the effect of external confounding variables which might have affected the results. In spite of the lack of individual randomisation, a comparison of these factors between the groups showed that the groups were fairly homogeneous before the intervention. The results show that the only variable affecting comprehension was the amount of reading time, with significant effects for students who read for more than two hours per day for pleasure. It is unlikely that these factors affected the results of the study because the variables were not statistically significantly different between the groups before the intervention. Further comparisons after controlling for these confounding factors failed to reveal a difference between the groups.

7.3. Discussion of MAI and verbal reports

Whereas the reading comprehension tests were designed to measure comprehension directly, the MAI and verbal reports were intended to measure the self-reported and observed utilisation of reading strategies. The answers from the self-reported MAI were first treated as interval data. The results from the analyses of the aggregated scores indicated that there existed no differences between the groups at baseline. After the intervention, the students in the strategy group reported higher strategy use than those in the control group. Also, the gain scores of the students in the strategy group were higher than those of the students in the control group.

The pre-test, post-test, and gain scores for each strategy were also compared between the groups. The results from this comparison of items clustered by strategy showed no difference between the groups for any strategy at baseline. However, after the intervention (i.e. at the same time as the comprehension immediate post-test took

place), the results indicated a difference in the perceived use of the following five strategies: 1) pre-reading strategies and goal setting, 2) the use of prior knowledge, 3) making inferences, 4) evaluating, and 5) strategy awareness; with the students in the strategy group reporting a higher use of these five strategies. Similarly, the comparison of the gain scores in each strategy indicated a statistically significant difference between the groups with regard to the following six strategies: 1) pre-reading strategies and goal setting, 2) reading forwards and backwards, 3) use of context, 4) monitoring and supervising, 5) evaluating, and 6) strategy awareness. Again, the students in the strategy group made higher gains.

Another set of analyses was carried out where the responses were treated as a categorical variable. The results of these analyses yielded similar results to when the data were treated as interval data. Both groups reported similar strategy-use at baseline except for four items (Q1, Q13, Q26, and Q27), and the control group reported higher use for three of these items. After the intervention, however, this difference increased to 12 items, with the students in the strategy group reporting higher strategy use. The results from the McNemar test, which tested the change in reported strategy-use over time (i.e. in the pre-test and post-test), showed changes in the percentages of students reporting strategy-use with regard to 20 items in the strategy group, but only with regard to three items in the control group. The students in the control group, however, never outperformed those in the strategy group. It is worth noting that these items were clustered in those strategies with regard to which the strategy group outperformed the control group in the analysis of the clustered items treated as interval data. Hence, the results from both analyses (i.e. as interval data, and as categorical data) yielded similar results, suggesting that the findings are robust to the choice of analytical approach. In sum, the students who participated in the intervention reported a higher frequency of strategy use than those in the control group.

The attempt to group items into strategies using an exploratory factor analysis did not show any conclusive clustering of items that was stable over time. A possible explanation for this is that the analysis uses as a basis the students' responses to cluster the items. Because the students' responses differed for the pretest and post-test, respectively, the analysis yielded different clusters each time the different datasets (pre and post) were analysed. Factor analysis is a widely used technique for validating questionnaires in order to verify the construct that is being measured (e.g. Gass, Demsky and Cameron Martin, 1998). A drawback of this method is that it is normally used once in order to validate instruments (i.e. questionnaires). To the best

of my knowledge, the process is rarely repeated to verify whether it measures the same construct when applied to a different sample (e.g. a different set of students, or a different point in time).

The results from the verbal reports showed that the students in both groups used strategies before the intervention. After the intervention, the number of strategies reported to be used by the students in the strategy group increased as well as the number of students using the strategies. The results indicate that the students in both groups used three different strategies before the intervention. While this number remained the same for the control group after the intervention, the number of strategies used by the students in the strategy group increased to ten. Five of these strategies were frequently used: quickly scanning, scanning to locate, reviewing and imagining, rereading to understand, and reading forwards and backwards. It is worth noting that these strategies were taught at the beginning of the intervention, which allowed for plenty of practice (see Section 4.6.2). The proportion of students using strategies remained similar for the control group (6-8 of the 19 students) but not for the strategy group (9-29 of the 40 students).

A comparison of the detection of incongruences at pre-test and post-test across the groups helped to measure the monitoring and supervision of the students' comprehension of the texts. The results show that the students in both groups improved in terms of the number of incongruences detected in the post-test; larger gains were achieved by the strategy group.

Although the results from both of these instruments (i.e. MAI and verbal reports) showed an increment in the use of strategies, this increment was not uniform across the strategies. Some strategies were mentioned more frequently than others, some strategies were not reported as being used at all, and the reported usage of the other strategies did not increase over time. Three reasons might explain these results. First, the effect of the intervention (i.e. the direct teaching of reading strategies) might not be the only variable that contributes to comprehension improvement. Paris, Lipson and Wixson (1983), who discussed the three types of knowledge (i.e. declarative, procedural and conditional), suggested that learning this type of knowledge is insufficient to improve comprehension. The students also need to be motivated and see that the goals of the strategies have personal relevance and meaning for them. This was explained to the students at the beginning of the project, in order to tackle this problem, as well as to motivate them and give them a purpose during the intervention. However, if the students were not intrinsically motivated, this might had caused a lack of increased usage of particular strategies.

Second, it may be that the students did not use all of the strategies all the time as intended during the design of the intervention. Although the students were taught all of the strategies, it remains their personal choice which strategies to use, and may depend on which strategies they anticipate will help them to solve their comprehension problems (Garner and Kraus, 1981).

Finally, this variation might be due to the time allowed to practice each strategy and automatise the skill. As mentioned in the methodology chapter (Chapter 4), the strategies were introduced on a weekly basis. The results showed greater changes in the use of the strategies that were taught at the beginning of the intervention (see Section 4.6.2). For example, S1 (pre-reading and goal setting), S2 (use of prior knowledge), S4 (making inferences), S7 (reading forwards and backwards), S9 (use of context), S14 (monitoring and supervising), S15 (evaluating), and S16 (strategy awareness) were taught at an early stage, and were amongst those that were reported to be used most frequently. This finding suggests that these strategies might have been better acquired because a) the students were exposed to them for a longer period of time, allowing automatisation, or b) because they were the most 'novel' at the start of the intervention. In contrast, the changes in the use of the strategies learnt at the end of the intervention were not significant because the students might not have being exposed for long enough to them and/or the intervention's novelty was reduced, thereby reducing its impact.

7.4. Discussion of the combined instruments

Further analyses were carried out to test whether there was a relationship between the students' perceptions of their strategy use and their reading comprehension scores. The results failed to indicate a correlation between the results for the reading comprehension test and the reported use of strategies (MAI). This lack of correlation between the scores for the reading comprehension test and the MAI could be explained by different factors. First, the extra reading practice might have helped to improve the students' comprehension solely through practice, but without the use of strategies. Second, the reading comprehension test, the MAI, and the verbal reports are measured using different scales. Due to the nature of the scales (e.g. one strategy is measured by several items in the MAI), it is possible that one scale (e.g. that of the reading comprehension test) is more sensitive' to improvement. Third, the students got used to the type of test; however, this threat was controlled by the different versions. Finally, the students might misreport their use of strategies due to

automatisation. When students have automatised and acquired a strategy, it becomes, arguably, an unconscious process (i.e. an automatised skill) (Mitchell and Myles, 2004), which is unavailable for conscious reflection and lies below the threshold of awareness. Thus, the reported strategy use might not reflect the actual use of all the strategies because the students might be unaware of the use and frequency of their skills. Both the questionnaire and the think-alouds rely on information provided by the students themselves, and so both are vulnerable to this threat to their validity. It is also important to emphasise the fact that successful learners apply strategies according to their own style (see the learning strategies for reference). Therefore the students might not have used all of the strategies in the way they were taught, and the taught strategies might have evolved according to the students' specific learning styles and experiences.

7.5. Comparison with previous studies

One of the motivations for this study was to test whether previous findings of the effectiveness of the direct teaching of reading comprehension strategies in English speakers would carry over to the Mexican setting. It is therefore useful to compare the findings of this study to those reported in the literature. Given the sizeable body of literature on this topic, the comparison focuses on seven studies carried out since 2000. These studies are most closely related to the research reported here and influenced the design of the current study (see Chapter 3).

The results from this study showed an initial improvement in reading comprehension for students in the strategy group at the end of the first stage of the intervention. This finding is in line with Boulware-Gooden et al. (2007); Cubukcu (2008); Inchausti de Jou and Sperb (2009); Wilawan (2007) and Madariaga Orbea and Martínez Villabeitia (2010), who reported initial gains immediately after their interventions. Thomas and Barksdale-Ladd (2000) did not measure reading comprehension. This study also found that students who were part of the strategy group and had been taught reading comprehension strategies maintained their initial gains over a four month period. None of the studies directly compared gains over a prolonged period of time, although Houtveen and Grift (2007) compared levels (i.e. post-test not adjusted for pre-test reading comprehension) after a one-year follow up and observed higher reading comprehension in the strategy group. Similar to this study, both Wilawan (2007) and Madariaga Orbea and Martínez Villabeitia (2010) reported increases in comprehension for the control groups at the immediate post-test, although in their

case the effect was not statistically significant. But, as both studies lacked a delayed post-test, no statements about maintenance or regression to the mean can be made.

This study found an increase in student self-reported use of strategies using a metacognitive assessment instrument and verbal reports. However, there was no statistically significant correlation between strategy use as assessed by the MAI and the verbal reports. No other study measured strategy use using two different instruments. Hence, no findings have been reported that confirm or reject the observed lack of correlation.

Generally speaking, comparisons with existing studies are made difficult by substantial differences in the methodology employed. For example, all studies used different assessment instruments so that findings are not directly comparable. Similarly, the studies varied in both the length of the intervention and the follow-up time.

8. Conclusions

It has been argued in educational research that teaching reading strategies improves students' comprehension. This study was designed to build on previous research by testing whether the direct teaching of reading comprehension strategies can be effectively applied to a different context. Three aspects of the study added to the previous research: 1) the direct teaching of reading strategies to help students to understand unfamiliar texts written in the Spanish language, 2) the duration of the intervention - a long intervention and delayed post tests in order to observe the maintenance of the results, 3) the inclusion of teachers to deliver the intervention, and 4) the combination of comprehension tests with self-reported instruments.

This chapter is divided into five parts. First, a brief summary of the study is provided. Second, the main findings of the study are summarised. Third, the limitations of the study are discussed as well as the process whereby the original design was adapted to fit in with the school regulations. Fourth, the implications for teaching and research are outlined. Finally, suggestions for further research are discussed.

8.1. Summary of the project, aims and methods

Research has shown that the direct teaching of reading strategies can improve students' reading comprehension. However, the effectiveness of this approach may depend on the specific design of the strategy instruction. Previous research suggests that teaching groups of strategies results in the better learning of reading comprehension strategies than teaching individual strategies in isolation (see Chapter 3). This study was designed to deliver an intervention that included the direct teaching of 16 of the most widely-researched reading comprehension strategies. The teaching techniques used in previous strategy teaching approaches, such as modelling, scaffolding, and cooperative learning, were included. The instruction aimed to teach what a strategy is and how to use it as well as when to use it and why (i.e. metacognitive information). The study tested two main hypotheses: a) the direct

teaching of reading comprehension strategies helps to improve 15 year-old Mexican students' comprehension of unfamiliar Spanish texts, and b) the students' awareness of strategy-use and their actual use of the strategies would change as a result of the intervention.

A sample of 177 secondary students in Mexico City participated in the study. The sample was divided into two groups: a strategy group and a control group. Both groups read and worked with the same materials. This was intended to reduce threats to the validity brought about by differences in the teaching environment that were unrelated to strategy teaching. Reading comprehension strategies (cognitive and metacognitive) were directly taught to the students in the strategy group whereas the students in the control group did not receive any explanation about the use of reading strategies. Both groups performed exactly the same tasks. The strategies were practised in collaborative groups following a scaffolding model.

The intervention was divided into three stages. The first stage was led by the researcher who introduced the strategies to the learners and oversaw controlled activities. The second stage was led by two teachers who worked with more independent tasks. At this point, the immediate post-tests were administered. During the third stage, the students were encouraged to practise the strategies, and no further teaching of strategies or reading tasks was provided. At the end of the third stage, the delayed post-tests were administered.

Three instruments were used to collect data as outcome measures: a reading comprehension test, a metacognitive assessment questionnaire, and verbal reports. The students' basic reading skills were tested before the intervention to establish whether the students possessed adequate decoding skills. The students also completed a background questionnaire which collected data on socio-economic background as well as reading activities and opportunities at home.

8.2. Summary of the findings

The students' comprehension was measured via a reading comprehension test, which was developed based on PISA texts and PISA-like questions. Three equivalent versions were used to measure the students' comprehension before (pre-test), immediately after (immediate post-test), and four months after the intervention (delayed post-test). Changes in comprehension were measured as follows: 1) changes between the pre-test and immediate post-test, 2) changes between the immediate post-test and delayed post-test, and 3) changes between the pre-test and delayed post-test.

The results from the test immediately after the intervention showed a statistically significant increment in the students' comprehension in both groups. Although such improvement could arguably be explained by the effect of the inclusion of a novel activity in the students' curriculum and a test effect, the results showed a larger effect size in the gains made by the strategy group. The results of the delayed post-test showed that there was a statistically significant drop in the control group's scores but not those of the strategy group. Hence, the students in the strategy group maintained the majority of gains four months after the end of the intervention, whereas those in the control group returned to the pre-intervention levels. These results suggest that, despite the fact that both groups might have been affected by the novelty and/or test effect at the immediate post-test, the teaching variable (i.e. strategies) helped the students in the strategy group to maintain the gains they had made at the delayed post-test; any novelty or test effect would have affected both groups similarly. To confirm these results, the gains between the pre- and delayed post-tests were compared. The results showed that the comprehension of both groups improved but that the gains in the control group were not statistically significant. The comprehension gains in the strategy group were statistically significant and showed comprehension improvement across time.

Due to the practical difficulty of randomising students at the individual level, a difference at baseline was observed between the students in both groups. This difference was not statistically significant. However, to ensure that the results were robust, an analysis of covariance was carried out which included the pre-test score as a covariate. The results of a paired comparison after controlling for this difference at baseline confirmed the maintenance of comprehension test gains in the strategy group at the delayed post-test as well as the non-maintenance of these by the control group. The inclusion of other potentially confounding factors, such as parents' socio-economic status, the reading opportunities at home, or class attendance during the intervention, did not alter this finding qualitatively.

The results from the self-report strategy questionnaire (MAI) showed changes in the students' perception of their strategy use after the intervention. An analysis of the overall scores showed that both groups possessed similar perceptions before the intervention; however, the strategy group increased their reported strategy-use significantly more than did the control group. A comparison of the perceptions of specific strategy-use showed that the students in both groups reported a more frequent use of five strategies after the intervention: pre-reading strategies, the use of prior knowledge, making inferences, evaluating, and strategy awareness. However,

the students in the strategy group constantly reported a higher use of strategies.

A comparison of the gain scores in the perceived use of strategies indicated that statistically significant differences exist between the groups with regard to the use of six strategies: pre-reading strategies, reading forwards and backwards, use of context, monitoring and supervising, evaluating, and an awareness of strategies. The results showed that the strategy group always enjoyed higher gains.

The responses to the self-report strategy questionnaire were also analysed as binary responses. The frequencies of the responses before and after the intervention were compared for each of the items on the questionnaire. Both groups showed similar strategy use at baseline (i.e. before the intervention) except for four items, whereby the students in the control group reported a higher usage than those in the strategy group. After the intervention, the results showed differences between the groups with regard to 12 items, whereby the students in the strategy group reporting a higher use of the strategy than those in the control group. A comparison of the changes over time per item showed that the reported strategy use was higher for 20 items in the strategy group and for 3 items in the control group.

The items on the questionnaire were tested to verify whether they were clustered as suggested in the previous literature (e.g. Mokhtari and Reichard, 2002). The results from the exploratory factor analysis suggested that the items would cluster into 15 factors. However, a clear cluster was not observed after the intervention. The analysis of the questionnaires after the intervention showed a different cluster than that existing before the intervention. The lack of clear, consistent clustering suggested a methodological problem with this test, which is discussed as a limitation of the study (see below).

To address concerns that self-report questionnaires fail to reflect accurately actual strategy use, the online use of strategies was measured by students performing verbal reports (i.e. think aloud protocols) during reading comprehension. The analysis of the transcripts showed that the students in the strategy group used or mentioned more strategies during reading tasks after the intervention. Some strategies were more frequently reported, possibly due to the fact that some strategies were practised for longer than others. After the initial introduction of strategies contained in the prompt used during the intervention (see Chapter 4), the strategies were introduced on a weekly basis. The results suggest that those strategies which were practised for longer were better automatized by the students than those that were only introduced at the end of the intervention.

The students' basic reading skills were tested at the beginning of the study. The

results of this test showed that the students possessed good decoding skills. The results from the vocabulary test showed that the students were better at obtaining the meaning of words placed in a context than that of words presented to them in isolation. This might explain why the strategy of reading forwards and backwards' was one of those that were used most frequently. The students might use this strategy in order to understand the meaning of vocabulary.

In addition to the reporting of the cognitive processes in the think-alouds, the detection of incongruences in the texts was also measured. The results showed no statistically significant difference between the numbers of incongruences detected before and after the intervention by the students in the control group. Although the students in the strategy group detected more incongruences after the intervention, this difference was not statistically significant. A comparison of the gain scores in the detection of incongruences for both groups showed larger gains for the students in the strategy group. This might mean that the students in the strategy group were better at monitoring their understanding of unfamiliar texts and/or were paying more attention to the micro details of the text.

It was hypothesised that an improvement of comprehension might be positively correlated with the number of strategies used by the students. To test this hypothesis, the gain scores from the reading comprehension test and the self-reported questionnaire were compared. The results however failed to show a correlation between the frequency of strategy-use and the improvement in comprehension.

These findings helped to answer the five research questions:

- Q1: Does the direct teaching of cognitive and metacognitive reading comprehension strategies help 15-year old Mexican students to improve their reading comprehension of a Spanish text?

  Response Q1: Yes. Although there was an immediate improvement in both groups, the results showed that comprehension gains were maintained only in the strategy group.

- Q2: Does the intervention raise the students' self-reported awareness of reading strategies and the reported frequency of use?

  Response Q2: Yes. The students in the strategy group reported a higher use of reading comprehension strategies after the intervention than their peers in the control group.

- Q3: Does the intervention increase the use of reading comprehension strategies

when reading unfamiliar texts?

Response Q3: Yes. The results of the verbal protocols revealed a higher strategy use among students in the strategy group compared to those in the control group.

- Q4: Is there a relationship between the reported use of strategies and performance in the reading comprehension test?

  Response Q4: No. No statistically significant association between reported use and comprehension was found at the immediate or delayed post-test.

- Q5: Do other factors, such as the indicators of socio-economic status, baseline reading comprehension scores, or word recognition, influence the effectiveness of the intervention with regard to reading comprehension.

  Response Q5: Yes. The baseline scores positively affected reading comprehension as well as the time spent reading for pleasure. The results also showed that those students who read for more than two hours per day were more likely to obtain higher reading comprehension scores. However, even when all potential confounding variables were taken into account, the strategy group still maintained their gains at the delayed post-test, whereas the control group did not. The estimated gains were similar, whether the potential confounders were controlled for or not.

8.3. Limitations of the study

Designing and carrying out an experiment in a controlled experimental environment, such as a laboratory, allows researchers to isolate the effect of the variable of interest (e.g. an intervention)(C. Torgerson and D. Torgerson, 2003; D. Torgerson and C. Torgerson, 2008). However, true randomised controlled experiments are rarely possible in educational research, which increasingly relies on interventions carried out in the classroom (Moore, A. Graham and Diamond, 2003). The alternative, carrying out a quasi-experiment in a real school context where entire classes are randomly allocated to the treatment and the control, tries to mimic the effect of a truly random allocation but cannot ensure that external factors are equally distributed across the treatment and control groups. These external factors may affect the variable of interest and bias the estimate of the effectiveness of the intervention under study. Despite all of the efforts to maintain conditions as similar as possible in both groups, there were some limitations associated with the characteristics of this study.

8.3.1. Lack of randomisation

First, the school logistics impeded the randomisation of the students at the individual level. Individual randomisation might have provided a homogeneous distribution of the students' characteristics, such as their age, pre-intervention reading comprehension level and strategy use, as well as their socio-economic status. Instead, cluster randomisation was carried out at the classroom level. However, due to the small number of clusters, it remains uncertain whether the groups were equally balanced. In fact, comparisons at the baseline showed differences between the groups across the different tests.

One way to address this problem is to create groups within classes and allocate students to them so that all students in one group are pair matched to students in the other group with respect to their observed characteristics (e.g. baseline scores). Although each class was divided into two subclasses during the first stage of the intervention, these subclasses were regrouped again during the second stage when the teachers took charge of the teaching and practice. Therefore, it was unfeasible to randomise them using matched selection. Also, because the researcher withdrew from actively teaching the students, there was a lack of trained staff to teach the alternative class at the same time as the class teacher was teaching one subclass.

Another approach to controlling for the lack of randomisation in quasi-experiments is to allocate each group to both the intervention and control conditions at different times, i.e. a 'switch-over' design (Cook and Campbell, 1979; Shadish, Cook and Campbell, 2002). The original design of the study reported here included a further stage of teaching the strategies to the control group over 2 months. However, the duration of this training was reduced to a couple of sessions due to a lack of time and resources. Also, no further testing was possible due to school restrictions and activities (i.e. end of year tests). If further teaching had been possible, it would have been interesting to compare the students' gains when taught with strategies and when engaging only in extensive reading.

8.3.2. Length of the intervention

A commonly-voiced criticism in the literature is the relatively short duration of many interventions (Pressley, 2000). They are usually delivered over short periods of time[1],

[1] For example, in a meta-analysis of 23 L2 studies which implemented reading strategy teaching, Norris and Ortega (2006) found that in more than half of the interventions the post-test is administered eight weeks or sooner after the pre-test.

with substantial changes made to the provision of teaching and study workload. Therefore, positive results from the interventions have been linked with a novelty effect. The difference between the activities undertaken during the study period and the activities normally carried out, or the novelty of having external staff in schools, is seen as a source of contamination in the interventions. Pressley (2000) therefore argues that researchers should aim to provide longer interventions where the effect can be measured over an academic year. Mitchell and Myles (2004) also argue that, for skills to be learnt, automatisation needs to be consolidated.

While it seems intuitively correct to favour longer interventions over shorter ones, this can impose several practical problems. First, the school administration may be unable or unwilling to grant access over a prolonged period of time. Second, research activities are costly and longer interventions are therefore likely to require more financial resources. Third, intervention activities are often subject to change according to the activities of the schools and the availability of resources (e.g. time and staff). After all, the school administration often considers research work as an extra activity that can be modified at any time to avoid interference with the school's activities. Therefore, a long intervention is not always possible without interfering with the normal activities of the classes, and shorter interventions are more likely to be conducted according to the pre-specified research protocol. Finally, long experimental interventions also raise ethical concerns. On the one hand, the students allocated to the control group do not receive an opportunity to benefit from the intervention (if indeed, it is beneficial). Therefore, withholding a successful treatment may be seen as unethical. On the other, the students participating in an intervention that implements e.g. a new way of teaching might suffer if the intervention proves to be unbeneficial or even harmful. Shorter interventions reduce the exposure to non-beneficial interventions and allow the students in the control group to benefit from effective interventions sooner.

In summary, the optimal length of an intervention is difficult to determine and researchers must be aware of the costs and benefits of staging a longer intervention. This study implemented a longer intervention (6 months in total) and the additional information collected at the delayed follow-up made it possible to ascertain the maintenance effect in the strategy group and the regression to baseline in the control group. In retrospect, it was felt that the benefits of a longer intervention outweighed any potential costs or risks.

8.3.3. Lack of a passive control group

The effects of an active control group were previously discussed (see Section 7.2). The results of this study showed how the activities planned for the control group might have affected their performance. One way of dealing with the improvements in the control group due to the planned activities is the inclusion of a passive control group. This control group should not receive any extra teaching, i.e. they should receive normal teaching, and should only take the tests. This would make it possible to control for the effects of planned activities on the active control group and to investigate the joint effect of strategy intervention and a change in reading activities. Unfortunately, it was impossible to include a passive control group due to the school's regulations; the cohort needed to be given the same activities. It was impossible not to deliver the new workshop to a class. However, as pointed out earlier, the primary interest of this study was the effect of teaching reading strategies rather than the improvements in comprehension due to extra reading activities. Hence, the comparison of the strategy group and active control group is the most relevant basis for inference.

8.3.4. Teacher commitment

It was previously argued that a common factor in high achieving countries is the in-service training delivered to teachers (OECD, 2006). Teaching training for secondary teachers is not provided with enough frequency nor sufficiently targeted to the teachers' needs, so Dillon et al. (2011) suggest providing teachers with school-embedded professional learning opportunities. This study provided two teachers with an opportunity to acquire new teaching techniques. However, one of the teachers decided not to deliver the lessons as requested despite having received support and a gentle reminder from the principal of the school. This lack of commitment reduced the practice time during the second stage of the project, and could have reduced the chance of observing the effects of strategy training amongst the students in the strategy group. These students practised the strategies once a week, rather than twice, as originally planned. This lack of practice might be one of the reasons why some of the strategies were more automatized than others. It also shows the type of problems faced when implementing new activities in schools and the reliance of researchers on the commitment of the teachers and the school administration.

Unfortunately, it was impossible to interview the teachers in order to study their perceptions of the intervention and the techniques used. The teacher who delivered

the second stage of the intervention reported using the strategies with her students in a different school outside the main study and witnessing positive results. Interviewing the teachers might have provided insights into the feasibility of adopting and adapting the activities to suit their curriculum.

8.3.5. Factor analysis

An interesting finding was that factor analysis found neither significant nor reliable clusters of items in the MAI, and also that the clusters failed to match those previously reported in the literature. This finding made it difficult to validate the results found here against other studies, as the analysis in these was performed using a different item scaling. Also, the clusters found in the MAI pre-test were different from those found in the MAI post-test. This fact made it difficult to compare whether the students' perceptions of their strategy use had changed; the strategy clusters themselves had changed. Factor analysis is a technique used frequently by researchers to ascertain the construct validity of questionnaires in specific populations (Davis, 1944; Rattray and Jones, 2007; Kaiser, 1960). However, to the best of my knowledge, they have not been validated in different populations (e.g. with different samples, over time). The use of cluster-based questionnaires as a between-subjects factor - to measure change over time in clusters of items - requires further study.

8.4. Implications for teaching

Strategy instruction has been regarded as a difficult activity to include in classroom teaching (Pearson, 2008). This study demonstrates its feasibility in a real world context. The results indicated that the students, who received direct teaching of reading comprehension strategies, improved and maintained their reading comprehension and increased their reported and observed use of strategies. Although not all the strategies were positively acquired, the students' awareness of them improved. It remains unclear how often the strategies should be taught to ensure their appropriate acquisition or for how long the students should be encouraged to practise them. Moreover, it is unclear how to demonstrate that students use the correct strategy.

The lack of correlation between the reading comprehension test, the MAI and the verbal reports suggests that assessment should be carried out with more than one outcome measure to provide reliable results. By relying solely on one outcome measure, the teachers might not be assessing correctly what the students have learnt.

Complementing assessment methods provides a wider and more complete view of the problems. Self-reported instruments might be open to misinterpretation or misreport by the students.

The involvement of the teachers in the project showed that it is possible to train teachers to provide strategy instruction as well as to train students consciously to evaluate and monitor their comprehension. The instruction however was carried out with teachers who are familiar with literacy skills. It would be worth pursuing this work with teachers delivering lessons in different content classes (e.g. science, humanities) to encourage and practise the use of strategies specific to other areas. Then, the students would be more frequently in contact with the strategies and so might realise that different texts require different strategies. They would also have the opportunity to practise the strategies in different contexts and to adapt them to their needs.

8.5. Implications for reading theories

Reading theories assume that automatisation of lower-level processes and the use of reading strategies to solve comprehension problems will result in comprehension of texts. The findings of this study suggest that the knowledge of strategies in itself is not sufficient to solve comprehension problems. The reader's ability to identify problems and the selection of the correct strategy should also be considered. The findings showed that despite good word decoding and recognition, students were unaware of existing problems. Metacognition is not normally integrated in reading theories although it has played a critical role in comprehension in the present study. This suggests that metacognition should form a more integral part of the theory of reading as it allows the reader to identify potential comprehension problems and react accordingly. In addition, the verbal protocols showed that even when students were provided with a range of strategies, it was not always the case that they would choose the correct one. This contrasts to the usual theories of reading that assume that, once students are provided with the necessary strategies, they will choose the strategies that are most appropriate to resolve their comprehension problem.

8.6. Implications for learning theories

The results of this study also have implications for general learning theories, especially skill acquisition theory. The findings of the MAI study suggest that the self-reported awareness of strategy use, observed at immediate post-test, was not yet proceduralised and so was not related with actual use during reading. However, over time, and with the extra practice provided by the teachers, this declarative and cognitive knowledge had the opportunity to become proceduralised and automatised, which may have contributed to the maintained comprehension gains. In contrast, the comparison group could not draw on any knowledge of strategy use. This emphasises the importance of practice and maturation of new knowledge in the acquisition of a new skill. The study also helped to emphasise the importance of modelling for less capable readers to understand the strategy use process; a concept that is closely related to the idea of the 'Zone of Proximal Development' and more generally to social interactions as emphasised by constructivism.

8.7. Implications for applied research

The results from the study showed the feasibility of explicitly teaching reading comprehension strategies in order to help the students to comprehend Spanish texts. An important feature of the intervention was its delivery in a real context, which highlighted the difficulty of conducting research in a school context. An implication of these findings is that the issues of randomization, time constraints, and the poor participation of teachers should be taken into account when designing studies. Although problems in these areas might arise, careful planning and adequate statistical techniques can help to control their effects. Carrying out research in ecological contexts makes it possible to study the desired population behaving under normal conditions, thereby providing more robust and policy relevant results. In this sense, this study constitutes a pragmatic trial which evaluates the effectiveness of teaching reading comprehension strategies within existing classrooms and under participation of teachers who provided (part of) the intervention.

As noted above, there is a need for further methodological investigation into the use of questionnaires with multiple items per construct. This study found that different clusters were observed at different testing points, calling into question the validity of previous research which has used these tools and has claimed to have validated them using factor analysis (Mokhtari and Reichard, 2002). To address

this problem, the current study used both single item analysis and also developed intuitively appealing clusters based on the previous literature. However, it would be more satisfactory to have clusters of items that reliably cluster together across multiple questionnaire responses by the same participants.

The rationale of this study arose from the poor results of Mexican students in the area of reading in PISA tests. The results obtained from the study suggest a way of addressing this problem, namely the direct teaching of reading strategies.

8.8. Further research

The results and limitations faced during the development of this study have resulted in a personal agenda for future research. There are several ways in which the research presented here might be expanded. First, future research should address the limitations of the present project and explore the robustness of the findings by incorporating a larger sample of different schools into a cluster RCT. It would be especially important to include two control groups - one active and one passive - to control for the effects of increased teaching activities in the active control group.

Second, more research is needed to establish the optimal length of an intervention. Longer interventions would allow studying the maintenance effect and ensure that the observed effect of teaching reading comprehension strategies is not due to a novelty effect. Also, by employing a switch-over design, researchers could compare the comprehension gains of a group when treated as a control group against its comprehension gains when treated as an experimental group. But, as stated before, longer interventions generally require more resources, raise ethical concerns, are likely to lose teacher commitment over time, and are difficult to integrate into the daily operation of a school. The costs of running a longer intervention must therefore be balanced against the potential gains.

Third, I would like to pursue further research on the feasibility of using factor analysis to analyse questionnaires with a pre- and post-test design. The research presented in this study raised questions about the clustering of items into strategies in existing instruments, and how stable this clustering is when the instrument is applied to different cohorts over different time points.

Fourth, this study employed both subjective and more objective measures of reading strategy use and reading comprehension, and found little correlation. However, many studies rely on only one instrument and it is therefore not possible to judge how robust their findings are and whether their results depend on the way in which

success was measured. A systematic review of this methodological issue was deemed outside the scope of this study; however such a study would be helpful and could make an interesting contribution to the educational research literature.

Fifth, and related, training teachers to deliver the intervention independently and over the course of an academic year would make it possible to deliver longer interventions and ensure that the results are not contaminated by the addition of a new, external teacher variable (i.e. the researcher). This may greatly enhance the external validity of the study and contribute to the professional development of teachers more generally. However, such efforts should go hand in hand with further studies about the teachers' perceptions regarding the direct teaching of reading comprehension strategies. As mentioned previously, one of the two teachers in this study lacked commitment during the second stage of the intervention, and one would expect some opposition from teachers when implementing changes to their teaching approach. Therefore, it is important to understand why teachers may be unwilling to participate in a study of reading comprehension strategies and develop reading programmes that suit not only the students but also the teachers' needs.

Sixth, this study tested the effectiveness of teaching reading comprehension strategies in the context of language classes. Future studies should investigate the usefulness of teaching reading comprehension strategies by teachers of other subject areas. This is likely to involve some tailoring of strategies and materials to specific subjects. Currently, the research in this area is scarce and focussed mainly on the strategies used by history and science teachers (Bean, 2000). However, given the importance of reading and comprehension for academic success, one may hypothesise that reading strategies can play an important part in other subject areas as well.

Finally, there is scope to investigate further which factors explain the country differences in reading performance as reported by PISA. For example, it would be interesting to compare the prevalence and use of reading strategies in different countries, and correlate this with the PISA results.

A. Appendix for Chapter 2

A.1. Comparison of educational stages (Mexico and UK)

Mexico			UK equivalent
Description	Stage	Age	
Kindergarten	1	4	Foundation Stage
	2	5	Key Stage 1
	3	6	
Primary School	1	6*	Key Stage 2
	2	7	
	3	8	
	4	9	
	5	10	
	6	11	Key Stage 3
Secondary School	1	12	
	2	13	
	3	14	Key Stage 4
High School	1	15	GCSE
	2	16	Sixth Form
	3	17	A Levels
Higher Education	4-5	$\geq$18	Undergraduate and postgraduate studies

(*) Due to a new law, which makes kindergarten attendace compulsory, students start school at different ages. Some students start aged 6 while others start at the age of 7.

A.2. Reading results from LLECE

Country	3rd grade	Country	4th grade
Cuba	343	Cuba	349
Argentina	263	Chile	286
Chile	259	Argentina	282
Brazil	256	Brazil	277
Venezuela	242	Colombia	265
Colombia	238	**Mexico**	**252**
Bolivia	232	Paraguay	251
Paraguay	229	Venezuela	249
Mexico	**224**	Honduras	238
Dominican Republic	220	Bolivia	233
Honduras	216	Dominican Republic	232

Reading in PERCE

Source: Martinez Rizo (2008)

Country	3rd grade	Country	4th grade
Cuba	637	Cuba	596
Uruguay	578	Costa Rica	563
Costa Rica	549	Chile	546
Mexico	**542**	Uruguay	542
Chile	517	**Mexico**	**530**
Argentina	513	Brazil	520
Brazil	499	Colombia	515
Colombia	493	Argentina	506
Paraguay	490	El Salvador	484
El Salvador	472	Peru	476
Peru	468	Nicaragua	473
Ecuador	460	Panama	472
Nicaragua	458	Paraguay	455
Panama	456	Guatemala	451
Guatemala	452	Ecuador	447
Dominican Republic	416	Dominican Republic	421

Reading in SERCE

Source: Martinez Rizo (2008)

B. Appendix for Chapter 3

B.1. Initial list of effective reading strategies obtained as part of literature review

	Strategy	Author(s)
1.	Alter expectations as text unfolds	R. Brown et al. (1996)
2.	Ask someone for help	R. Brown et al. (1996)
3.	Browsing or previewing	R. Brown et al. (1996)
4.	Clarifying confusions	R. Brown et al. (1996)
5.	Deciding on the difficulty of text	Cubukcu (2008)
6.	Determine consistency	Pressley, El-Dinary et al. (1992)
7.	Distinguishing new and old information	Cubukcu (2008)
8.	Draw inferences	Simpson and Nist (2000)
9.	Elaborate	Simpson and Nist (2000)
10.	Evaluate	Simpson and Nist (2000); Inchausti de Jou and Sperb (2009); Cubukcu (2008)
11.	Generate questions & Self-questioning (to check understanding)	R. Brown et al. (1996); Simpson and Nist (2000); Gillies and Khan (2009); Phakiti (2003); Houtveen and Grift (2007); Pressley, El-Dinary et al. (1992); Pressley (2000)
12.	How to deal with unknown words or phrases & Problem solving tactic for vocabulary & Inferring meaning of words	R. Brown et al. (1996); Pressley, El-Dinary et al. (1992); Cubukcu (2008)
13.	Identify text structure	Inchausti de Jou and Sperb (2009)
14.	Interpretations	R. Brown et al. (1996); Pressley, El-Dinary et al. (1992)
15.	Keep idea of paragraphs in mind	Inchausti de Jou and Sperb (2009)
16.	Looking back	R. Brown et al. (1996)
17.	Main idea & Identify important aspects of message	Wilawan (2007); Phakiti (2003)
18.	Mapping (look for descriptions, cause-effect, time sequence, construct maps)	Pressley, El-Dinary et al. (1992)

Continued on next page

continued from previous page

Strategy	Author(s)	
19.	Monitoring	R. Brown et al. (1996); Simpson and Nist (2000); Phakiti (2003); Houtveen and Grift (2007)
20.	Number paragraphs	Inchausti de Jou and Sperb (2009)
21.	Organising strategies (concept maps, network representations)	Simpson and Nist (2000)
22.	Plan	Simpson and Nist (2000)
23.	Predicting (predicting using titles, subheadings, summary, layouts) & Guessing the later topics (anticipate info)	R. Brown et al. (1996); Gillies and Khan (2009); Houtveen and Grift (2007); Pressley, El-Dinary et al. (1992); Cubukcu (2008)
24.	Read selectively (change speed)	Houtveen and Grift (2007)
25.	Reading goals	Cubukcu (2008)
26.	Relate information	Gillies and Khan (2009)
27.	Relate prior knowledge or personal experiences to text using background information, knowledge, prior knowledge activation	R. Brown et al. (1996); Houtveen and Grift (2007); Pressley, El-Dinary et al. (1992); Cubukcu (2008); Pressley (2000)
28.	Rereading	R. Brown et al. (1996); Inchausti de Jou and Sperb (2009)
29.	Revising prior questions	Cubukcu (2008)
30.	Searching according to goals	Cubukcu (2008)
31.	Select and transform ideas	Simpson and Nist (2000)
32.	Self regulate comprehension	Inchausti de Jou and Sperb (2009)
33.	Set a goal/purpose	R. Brown et al. (1996); Phakiti (2003)
34.	Skipping	R. Brown et al. (1996)
35.	Story grammar analysis	Pressley, El-Dinary et al. (1992); Pressley (2000)
36.	Substituting or guessing	R. Brown et al. (1996)
37.	Summarise or Retell	R. Brown et al. (1996); Simpson and Nist (2000); Gillies and Khan (2009); Pressley, El-Dinary et al. (1992); Pressley (2000)
38.	Take corrective action when failure	Phakiti (2003)
39.	Text structure analysis (webbing)	Pressley, El-Dinary et al. (1992)
40.	Think aloud	R. Brown et al. (1996); Pressley, El-Dinary et al. (1992)
41.	Think questions that the teacher would ask	Gillies and Khan (2009)
42.	Underline key words	Inchausti de Jou and Sperb (2009)
43.	Using strengths	Cubukcu (2008)
44.	Verifying	R. Brown et al. (1996)
45.	Visualizing & Constructing images	R. Brown et al. (1996); Pressley, El-Dinary et al. (1992); Pressley (2000)

B.2. Approaches to teaching reading comprehension strategies

Approach	Strategies taught	Observations
Reciprocal teaching	summarising; questioning; seeking clarification; predicting	There is a gradual release of responsibility
Transactional Strategies Instruction (TSI)	predicting; verifying; visualizing; relating prior knowledge; questioning; clarifying; making associations; summarising; monitoring; looking back	It teaches students when to use each of these strategies - involves a socially constructive meaning
Collaborative Strategic Reading (CSR)	preview the text before reading; click and clunk to monitor; getting the gist while reading; wrap up after reading	
Scaffolded Reading Experience (SRE)	activate prior knowledge; making predictions; questioning; discussions; drawing; graphic organisers	It involves teaching vocabulary during the lesson
Peer Assisted Learning Strategies (PALS)	retelling; summarising; predicting;	
Questioning the Author (QtA)	questioning	
Cognitive Academic Language Learning Approach (CALLA)	It teaches different strategies by applying the following steps: preparation, presentation, practicing, evaluation, applying (expansion)	

C. Appendix for Chapter 4

C.1. Words used in word identification test

High frequency	Medium frequency	Low frequency
2 and 3 syllables (40%)		
interés	desvío	llévame
amigo	acogedor	ladeado
paso	falsedad	serrana
derecho	castaño	marasmo
	aplacar	baremo
4 and 5 syllables (40%)		
actividad	terrateniente	alargamiento
movimiento	ofrecimiento	halagador
entender	aturdido	estupefacción
desarrollo	próximamente	mordacidad
	rudimentario	carterista
6 or more syllables (20%)		
N/A	desaparecido	sospechosamente
	imaginativo	intuitivamente
	acondicionado	telefónicamente
		evolucionista

C.2. Words used in vocabulary test

	Spanish	English
1.	cuchillo	knife
2.	paraguas	umbrella
3.	reloj	clock
4.	sombrero	hat
5.	bicicleta	bike
6.	clavo	nail
7.	abecedario	alphabet
8.	burro	donkey
9.	ladrón	thief
10.	juntar	join
11.	valiente	brave
12.	diamante	diamond
13.	apostar	to bet
14.	disparate	nonsense
15.	prevenir	to prevent
16.	contagioso	contagious
17.	molestia	discomfort
18.	fabula	fable
19.	peligroso	dangerous
20.	emigrar	migrate
21.	estrofa	verse
22.	recluir	seclude
23.	escarabajo	beetle
24.	espionaje	espionage
25.	campanario	clock tower
26.	rivalidad	rivalry
27.	reforma	reform
28.	impulsar	boost/push
29.	aflicción	affection
30.	demoler	demolish
31.	inminente	imminent
32.	dilatorio	dilatory

C.3. Word pairs used in similarities test

	Spanish	English
1.	rueda - pelota	wheel - ball
2.	vela - lámpara	candle - lamp
3.	camisa - sombrero	shirt - hat
4.	piano - guitarra	piano - guitar
5.	manzana - plátano	apple - banana
6.	cerveza - vino	beer - wine
7.	gato - ratón	cat - mouse
8.	codo - rodilla	elbow - knee
9.	teléfono - radio	phone - radio
10.	kilo - metro	kilo - meter
11.	enojo - alegría	anger - happiness
12.	tijeras - sartén	scissors - frying pan
13.	montaña - lago	mountain - lake
14.	libertad - justicia	freedom - justice
15.	primero - ultimo	first - last
16.	49 y 121	49 and 121
17.	sal - agua	salt - water

C.4. Word-level study

Nombre_____Grupo_____

	Año	Mes	Día
Fecha de Aplicación			
Fecha Nacimiento			
Edad			

SEMEJANZAS Descontinuar después de **3** fracasos consecutivos	Puntos 1,0
1. rueda – pelota	
2. vela – lámpara	
3. camisa – sombrero	
4. piano – guitarra	
5. manzana – plátano	
6. cerveza – vino	
7. gato – ratón	
8. codo – rodilla	
9. teléfono – radio	
10. kilo – metro	
11. enojo – alegría	
12. tijeras – sartén	
13. montaña – lago	
14. libertad – justicia	
15. primero – ultimo	
16. 49 y 121	
17. sal - agua	
*Si el niño da una respuesta de 1 punto, diga: "En que más se parecen los números 49 y 121?"	

VOCABULARIO Descontinuar después de **5** fracasos consecutivos	Puntos 2,1,0
1. cuchillo	
2. paraguas	
3. reloj	
4. sombrero	
5. bicicleta	
6. clavo	
7. abecedario	
8. burro	
9. ladrón	
10. juntar	
11. valiente	
12. diamante	
13. apostar	
14. disparate	
15. prevenir	
16. contagioso	
17. molestia	
18. fabula	
19. peligroso	
20. emigrar	
21. estrofa	
22. recluir	
23. escarabajo	
24. espionaje	
25. campanario	
26. rivalidad	
27. reforma	
28. impulsar	
29. aflicción	
30. demoler	
31. inminente	
32. dilatorio	
Total Max.=64	

C.5. Reading comprehension test - version A

THE UNIVERSITY of York

Nombre: _____ **Grupo:** _____

NL: _____

EDITORIAL

La tecnología crea la necesidad de nuevas leyes

La ciencia tiene una forma de rebasar a las leyes y a la ética. Eso sucedió dramáticamente en 1945 desde la perspectiva de la destrucción de la vida con la bomba atómica, y está sucediendo ahora desde la perspectiva de la creación de la vida con las técnicas para remediar la infertilidad humana.

La mayoría de nosotros nos alegramos con la familia Brown en Inglaterra cuando Louise, el primer bebé de probeta, nació. Y nos hemos maravillado con otros indicios - últimamente con los nacimientos de bebés sanos que alguna vez fueron embriones congelados en espera del momento adecuado para el implante de la futura madre.

Y es con respecto a dos embriones así congelados en Australia que se desencadenó una polémica discusión seguida de una tormenta de problemas legales y éticos. Los embriones estaban destinados a implantarse en Elsa Ríos, esposa de Mario Ríos. El implante previo de un embrión había fracasado, y los Ríos querían tener otra oportunidad de ser padres. Pero antes de tener la oportunidad de un segundo intento, los Ríos perecieron en un accidente aéreo.

¿Qué debía hacer el hospital australiano con los embriones congelados? ¿Podían ser implantados en alguien más? Había numerosas voluntarias. ¿Tenían los embriones derecho, de alguna manera, al considerable patrimonio de los Ríos?

¿O debían ser destruidos los embriones? Los Ríos, comprensiblemente, no habían previsto el futuro de los embriones.

Los australianos nombraron una comisión para estudiar el asunto. La semana pasada, la comisión presentó su informe. Los embriones debían ser descongelados, estableció el grupo, porque la donación de los embriones a otra persona requeriría el consentimiento de los "productores" y no se había dado el consentimiento.

El grupo estableció también que los embriones en el estado actual no eran seres vivos ni tenían derechos, y por lo tanto podían ser destruidos.

Los miembros de la comisión estaban conscientes de que se estaban adentrando, legal y éticamente, en terreno peligroso. Así que insistieron en que se esperara tres meses para que la opinión pública respondiera a la recomendación de la comisión. En caso de haber un clamor abrumador en contra de la destrucción de los embriones, la comisión reconsideraría.

Las parejas que ahora se inscriben en los programas de fertilización in vitro del hospital Reina Victoria en Sidney, tienen que especificar con anterioridad qué debe hacerse con los embriones en caso de que algo les suceda.

Esto asegura que no se volverá a dar una situación como la de los señores Ríos. ¿Pero qué hay de otras cuestiones más complicadas? En Francia, por ejemplo, una mujer tuvo que ir al tribunal para que se le permitiera tener un hijo con el esperma congelado de su difunto marido. ¿Cómo debe manejarse una petición tal? ¿Qué pasa si una madre sustituta rompe su contrato de gestación y se rehusa a dar el niño que había prometido gestar para alguna otra mujer?

Nuestra sociedad ha fracasado hasta el momento en establecer leyes que puedan hacerse valer para restringir el potencial destructivo de la energía atómica. Estamos cosechando los espeluznantes frutos de ese fracaso. Las posibilidades del mal uso de la capacidad de los científicos para acelerar o retrasar la procreación son múltiples. Es apremiante, por lo tanto, establecer límites éticos y legales antes de que vayamos demasiado lejos.

Continua al reverso de la hoja

1A

Consulta el artículo de la página anterior para responder las siguientes preguntas. Circula la respuesta correcta.

1.- Las parejas que se inscriben a programas de fertilización en el hospital de Sidney deben especificar el procedimiento a seguir con los embriones en caso de accidente. Esta medida pretende:
   A. Tener un mejor control de los embriones
   B. Tener registros más completos en el hospital
   C. Evitar un problema como el que se dio con los Ríos
   D. Apoyar a las parejas que no pueden tener hijos.

2.- ¿Cuál es el objetivo principal del autor?
   A. advertir
   B. divertir
   C. informar
   D. convencer

3.- Menciona dos ventajas de las técnicas para remediar la infertilidad humana.

   a)_____

   b)_____

4.- Subraya en el texto, la oración que explica que hicieron los australianos para decidir que se debería hacer con los embriones congelados, que pertenecían a una pareja que murió en un accidente aéreo.

2A

Consulta la tabla y resuelve las preguntas de la siguiente página.

Resultados del Programa México hacia el 2000 durante el Año Fiscal de 1998

Región Sur

	Guerrero	Veracruz	Chiapas	Oaxaca	Puebla	Michoacán	Tabasco	Campeche	Yucatán	Total
Crecer saludablemente										
Construcción de clínicas de salud	1	0	0	0	2	1	3	2	1	10
Capacitación de personal de la salud	100	89	0	110	325	237	143	75	80	1159
Niños que recibieron complementos alimenticios	2500	1900	0	3000	859	427	890	420	350	10346
Niños que recibieron ayuda financiera para tratamiento dental	100	70	0	120	85	62	58	50	0	545
Aprender										
Maestros capacitados durante una semana	1120	860	0	0	2000	975	430	250	380	6015
Cuadernos escolares adquiridos/donados	300	150	0	0	225	180	0	285	0	1140
Niños que recibieron becas	20	0	0	3	35	0	15	0	12	85
Libros escolares donados/adquiridos	0	0	0	130	170	260	238	154	95	1047
Pupitres escolares construidos/donados	200	120	0	0	450	348	360	130	0	1608
Salones de clase construidos permanentemente	20	0	0	5	0	10	8	11	2	56
Salones reparados	0	30	0	12	45	32	50	10	0	179
Adultos que fueron alfabetizados	300	120	0	150	0	40	95	35	25	765
Hábitat										
Letrinas o retretes excavados/fabricados	40	30	0	30	13	25	48	67	80	333
Casas con servicio de drenaje	250	180	0	42	450	390	230	176	82	1800
Pozos mejorados	15	0	0	0	0	8	23	18	5	69
Nuevos pozos (excavados)	0	1	0	2	0	5	4	8	3	23
Nuevas redes de agua potable	0	0	0	0	0	15	5	8	0	28
Redes de agua potable reparadas	1	0	0	1	0	0	2	0	1	5
Casas remozadas	115	30	0	20	0	15	35	60	10	285
Nuevas casas construidas	100	0	0	0	250	80	75	80	0	585
Centros comunitarios construidos o mejorados	5	2	0	1	0	2	8	0	1	19
Entrenamiento de líderes comunitarios >1 día	500	214	95	80	450	320	270	130	60	2119
Kilómetros de caminos reparados	15.6	20.8	0	12	10	0	21	14	0	93.4
Puentes construidos	4	0	0	1	0	0	1	2	0	8
Familias beneficiadas por el control de la erosión	0	350	0	0	150	0	232	0	80	812
Casas que ahora cuentan con servicios de electricidad	530	423	0	370	265	170	98	85	70	2011

Continua al reverso de la hoja

3A

La tabla de la página anterior forma parte de un informe publicado por el Programa México hacia el Siglo XXI, una organización de asistencia. En ésta se muestra alguna información sobre el trabajo del programa en una de sus regiones de operación (Región Sur). Consulta la tabla para responder las siguientes preguntas. **Circula** la respuesta correcta.

5.- De acuerdo con la tabla, ¿En qué estado se realizó mayor numero de actividades?
   A. Puebla
   B. Veracruz
   C. Tabasco
   D. Yucatán

6.- ¿Qué dice la tabla sobre el nivel de actividad del programa en Chiapas en el 2000, en comparación con su actividad en otros estados del país?
   A. El nivel de actividad fue comparativamente alto en Chiapas.
   B. El nivel de actividad fue comparativamente bajo en Chiapas.
   C. Fue más o menos el mismo que en otros estados de la región.
   D. Fue comparativamente alto en la categoría de hábitat y bajo en otras categorías.

7.- En el 2000 Chiapas fue uno de los estados más pobres del país. Tomando en cuenta este hecho y la información de la tabla, ¿Qué crees que podría explicar el nivel de las actividades del programa realizadas en Chiapas, en comparación con sus actividades en los otros estados?

_____

_____

8.- A continuación encontrarás algunas otras actividades de asistencia. Si estas actividades se añadieran a la tabla, ¿a qué categoría pertenecería cada una? Indica tu respuesta **cruzando** el recuadro correcto, a continuación de cada una de las actividades.

	Crecer saludablemente	Aprender	Hábitat
Teléfonos públicos instalados	☐	☐	☐
Niños vacunados	☐	☐	☐
Madres que recibieron información sobre la alimentación de infantes	☐	☐	☐
Plantas solares de generación de energía eléctrica construidas	☐	☐	☐

Mueren tres mil peces por falta de oxígeno en un lago de Aguascalientes

CLAUDIO BAÑUELOS, CORRESPONSAL

Aguascalientes, Ags., 6 de marzo. Más de 3 mil tilapias plateadas murieron durante la semana pasada por falta de oxígeno en el lago ubicado en el parque ecológico El Cedazo, informó el titular del Instituto del Medio Ambiente del Estado (IMAE), Juan Solorio Tlaseca.

4A

Indicó que el parque fue cerrado al público el fin de semana pasado para realizar labores de limpieza en el embalse donde se reportó la mortandad de peces, y se espera que el miércoles próximo reabra sus puertas al público.
Solorio Tlaseca rechazó que ese problema fuera originado por la presencia de aguas negras en el estanque, como se dijo extraoficialmente.
Sostuvo que la causa del desastre ecológico es que el lago tiene baja concentración de oxígeno, recurso que necesitan los organismos vivos del embalse, como el plancton, por lo que los peces no están teniendo las condiciones adecuadas para vivir.
Señaló que personal del IMAE y la Secretaría de Agricultura, Desarrollo Rural, Ganadería, Pesca y Alimentación (SAGARPA) están trabajando para sacar con redes los peces muertos, para que la descomposición de éstos no afecte al resto de la población que incluso también está siendo retirada para su observación en estanques .
"Aunque desafortunadamente, las condiciones de los peces vivos no son muy adecuadas para rescatarlos; si no se pueden rescatar, los vamos a tener que sacrificar", agregó.
El funcionario precisó que la mortandad de peces se presentó desde la segunda semana de enero, ya que aunado a la falta de oxigeno, las bajas temperaturas que se presentaron en esas fechas también afectaron a las tilapias.
"Después de los fríos se complicó la situación de las tilapias por la baja disponibilidad de oxígeno en el embalse, que necesita de nutrientes, como algas, necesarios para que los seres vivos puedan sobrevivir en condiciones más favorables", dijo.

Martes 7 de marzo del 2006
La Jornada.

Consulta la información para contestar las preguntas. Circula la respuesta correcta.

9.- ¿Qué es lo que pretende el autor en este texto?
A. Resaltar la necesidad que tienen los organismos vivos del embalse para sobrevivir.
B. Destacar la presencia de aguas negras en el estanque.
C. Informar sobre la muerte de tres mil peces por falta de oxigeno.
D. Informar sobre la intervención de las autoridades ante la muerte de tres mil peces.

10.- De acuerdo con el artículo, ¿Por qué las bajas temperaturas afectaron a las tilapias?

_____

_____

11.- Mira la siguiente oración ubicada casi al final del artículo (se presenta en dos partes):

"Aunque desafortunadamente, las condiciones de los peces vivos no son muy adecuadas para rescatarlos;…" ***(Primera parte)***

"… si no se pueden rescatar, los vamos a tener que sacrificar." ***(Segunda parte)***

¿Cuál es la relación entre la primera y segunda parte de la oración?

Continua al reverso de la hoja

5A

La segunda parte:

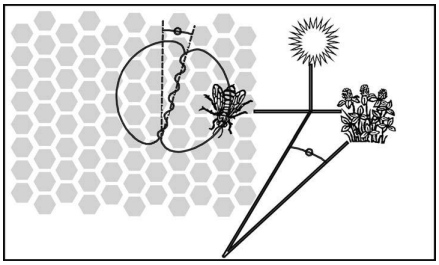
A. Contradice a la primera parte.
B. Confirma la primera parte.
C. Ilustra el problema descrito en la primera parte.
D. Da la solución al problema descrito en la primera parte.

LAS ABEJAS

RECOLECCIÓN DEL NÉCTAR

Las abejas fabrican miel para sobrevivir. Es su única fuente de alimentación. Si hay 60.000 abejas en una colmena, alrededor de una tercera parte está dedicada a la recolección del néctar que las abejas elaboradoras convertirán después en miel. Una pequeña parte de las abejas trabajan como exploradoras o buscadoras. Encuentran una fuente de néctar y luego vuelven a la colmena para comunicárselo a las otras abejas.

Las exploradoras comunican dónde está la fuente de néctar ejecutando una danza que transmite información sobre la dirección y la distancia que las abejas tendrán que recorrer. Durante esta danza la abeja sacude el abdomen de un lado a otro mientras describe círculos en forma de 8. La danza sigue el dibujo mostrado en el siguiente gráfico.



El gráfico muestra a una abeja bailando dentro de la colmena en la cara vertical del panal. Si la parte central del 8 apunta directamente hacia arriba, significa que las abejas encontrarán el alimento si vuelan directamente hacia el sol. Si la parte central del 8 apunta a la derecha, el alimento se encuentra a la derecha del sol.

La cantidad de tiempo durante el cual la abeja sacude el abdomen indica la distancia del alimento desde la colmena. Si el alimento está bastante cerca la abeja sacude el abdomen durante poco tiempo. Si está muy lejos, sacude el abdomen durante mucho tiempo.

PRODUCCIÓN DE LA MIEL

Cuando las abejas llegan a la colmena con el néctar, lo pasan a las abejas elaboradoras, quienes manipulan el néctar con sus mandíbulas, exponiéndolo al aire caliente y seco de la colmena. Recién recolectado, el néctar contiene azúcares y minerales mezclados con alrededor de un 80% de agua. Pasados de diez a veinte minutos, cuando gran parte del agua sobrante se ha evaporado, las abejas

6A

elaboradoras introducen el néctar dentro de una celda en el panal, donde la evaporación continúa. Tres días más tarde, la miel que está en las celdas contiene alrededor de un 20% de agua. En este momento, las abejas cubren las celdas con tapas que fabrican con cera.

En cada período determinado, las abejas de una colmena suelen recolectar néctar del mismo tipo de flor y de la misma zona. Algunas de las principales fuentes de néctar son los frutales, el trébol y los árboles en flor.

Consulta la información para contestar las preguntas. Circula la respuesta correcta.

12.- ¿Cuál es el propósito de la danza de la abeja?
- A. Celebrar que la producción de la miel ha sido un éxito.
- B. Indicar el tipo de planta que han encontrado las exploradoras.
- C. Celebrar el nacimiento de una nueva reina.
- D. Indicar donde han encontrado las exploradoras el alimento.

13.- Anota tres de las principales fuentes de néctar.

a) _____

b) _____

c) _____

14.- En la danza, ¿qué hace la abeja para mostrar la distancia existente entre el alimento y la colmena?

_____

_____

MI PRIMER AMOR

Tenía yo solamente trece años.
Ella era encantadora.
¡Qué digo encantadora! Era una de las mujeres más bonitas de París.
Pero de eso yo no me daba cuenta. Yo la *encontraba* bonita — ocurría que lo era extremadamente. Esto no era más que una coincidencia para mí...
...Tenía una sonrisa adorable y ojos acariciadores.
Y voy a preguntarme, ¿por qué la he amado?
...Soñaba a diario con ella.
¿Decírselo?
Antes mejor la muerte.
¿Entonces?
Probárselo.
Hacer economías durante toda la semana y cometer una locura el domingo siguiente. Hice estas economías y cometí esta locura. Ocho francos se convirtieron en un enorme ramo de violetas. ¡Era magnífico! Era el más bello ramo de violetas que se haya visto nunca. Me hacían falta las dos manos para llevarlo.
Mi plan: llegar a su casa a las dos en punto y solicitar verla.

Continua al reverso de la hoja

7A

La cosa no fue fácil. Estaba ocupada pero insistí. La camarera me condujo al gabinete donde ella se encontraba.
Se estaba peinando para salir. Entré con el corazón en un brinco.
—¡Hola, pequeño! ¿Para qué quieres verme?
No se había vuelto aún. No había visto todavía el ramo; no podía comprender lo que pasaba a sus espaldas.
—Para esto, señora.
Y le tendí mis ocho francos de violetas.
—¡Oh, qué bonitas!
Me pareció que la partida estaba ganada. Me había aproximado a ella, temblando. Cogió entre sus manos mi ramo como se coge la cabeza de un niño y lo llevó a su bello rostro como para besarlo.
—¡Y huelen bien!
Luego, añadió despidiéndome:
—Dale las gracias de mi parte a tu papá.

Sacha Guitry

15.- ¿Por cuánto tiempo ahorra el autor para poder comprar el ramo de violetas?
   A. 5 días
   B. un mes
   C. 7 días
   D. se lo pide a su padre

16.- El autor no quería decir a la señora lo que sentía por ella. Subraya en el texto la oración que lo demuestra?

17.- ¿Por qué creyó el autor que "La partida estaba ganada.."?

_____

_____

Aquí esta parte de la conversación entre dos personas que leyeron "Mi Primer Amor"

18.- Usa información de la lectura para mostrar como estas personas podrían justificar sus puntos de vista.

Persona 1

_____

_____

Persona 2

_____

_____

8A

C.6. Reading comprehension test - version B

THE UNIVERSITY of York

Nombre: _____ **Grupo:** _____

NL: _____

Bioética

Los avances en el campo de la biotecnología han generado una creciente demanda de supervisión y control de la manipulación de los fenómenos biológicos. El hallazgo de la estructura y funcionamiento del material genético suscitó en diversos sectores dudas acerca de los límites de la actuación humana en este campo.

La bioética se dedica al estudio de las implicaciones éticas y morales de los nuevos descubrimientos y avances biomédicos. Pretende establecer principios que hagan compatibles el desarrollo investigador con la ética, el respeto al medio y la protección del patrimonio genético del ser humano.

El concepto de bioética es muy antiguo. Entre los primeros centros de modernos estudios bioéticos cabe citar al Centro Hastings, fundado en 1969, y al Instituto de Ética Joseph y Rose Kennedy, creado en 1971.

A finales de los 90's, este tipo de instituciones había proliferado hasta contarse cientos de ellas en todo el mundo. Su espectro de actuación se refiere fundamentalmente a la ingeniería genética, aunque también se extiende a la investigación de fármacos y sus ensayos en animales y humanos, e incluso a temas polémicos como el aborto y la eutanasia.

La definición de un código ético resulta controvertida en el ámbito de la ingeniería genética. Al conocer los secretos de la transmisión de la información genética, los científicos pueden traducir el código genético de los seres vivos y llevar a cabo manipulaciones de diversa índole. Tal es el caso de las plantas transgénicas, como el maíz y la soja, que incluyen nuevos genes de resistencia a plagas o a fenómenos adversos como la sequía. Asimismo, en el campo de la experimentación animal se ha logrado una mayor productividad, consiguiéndose especies transgénicas de crecimiento más rápido.

No obstante, la mayor preocupación estriba en la utilización de la ingeniería genética en el ser humano. Por ejemplo, mediante la introducción de genes procedentes de grandes primates podrían obtenerse humanos transgénicos que desarrollaran una mayor fuerza muscular y que quedaran determinados genéticamente desde su nacimiento, para la realización de las labores más pesadas. La bioética se plantea hasta que punto y en qué condiciones se debe experimentar con los seres humanos. A tal respecto se refiere el protocolo firmado en 1998 en virtud del cual se prohibía la clonación de seres humanos.

Por otro lado, el Proyecto Genoma Humano pretende la lectura completa de la información genética contenida en los 46 cromosomas de la especie humana y la representación de todos sus genes, cuyo número se estima entre cincuenta mil y cien mil. Este proyecto tiene como primer objetivo la identificación y el aislamiento de los genes directamente implicados en enfermedades. La secuenciación total del genoma humano permitirá desarrollar una medicina preventiva que anticipe el tratamiento de enfermedades latentes, y constituye ya la base de la terapia génica, mediante la cual es posible reconstruir los cromosomas portadores de dolencias.

Sin embargo, frente a estas esperanzadoras perspectivas se abren nuevas interrogantes que afectan a la protección legal del individuo. Existe el riesgo de que, en determinadas circunstancias, las personas se vean obligadas a comunicar su identidad genética a las empresas demandantes de empleo o a las compañías aseguradoras. La presencia de un determinado indicador genético de ciertas enfermedades sería suficiente para incrementar el precio de las pólizas de seguros y podría impedir el logro de un puesto de trabajo.

1B

Continua al reverso de la hoja

Para asegurar la protección frente a actitudes discriminatorias, los planteamientos bioéticos deben plasmarse en textos legales. El propio Proyecto Genoma Humano tiene un grupo de trabajo para cuestiones éticas, legales y sociales encargado de regular las directrices del proyecto y controlar la difusión y aplicaciones del mismo.

Consulta el artículo 'Bioética' para responder a las siguientes preguntas. Circula la respuesta correcta.

1.- El Proyecto Genoma Humano pretende:
   A. Crear humanos transgénicos
   B. Crear un nuevo tipo de medicina génica
   C. Brindar una nueva forma de identificación a los individuos
   D. Aislar e identificar la información genética de la especie humana

2.- ¿Cuál es el objetivo principal del autor?
   A. advertir
   B. divertir
   C. informar
   D. convencer

3.- Menciona dos ventajas que nos proporciona el conocimiento de la información genética.

   a)_____

   b)_____

4.- A pesar de las múltiples ventajas de la biotecnología, hay un campo en donde su utilización es motivo de preocupación. Subraya en el texto la oración que expresa tal preocupación.

2B

Consulta la tabla y resuelve las preguntas de la siguiente página.

Resultados del Programa PLAN Internacional durante el durante el Año Fiscal de 1996

Región de África Oriental y del Sur

	Egipto	Etiopía	Kenia	Malawi	Sudan	Tanzania	Uganda	Zambia	Zimbabwe	Totales
Crecer saludablemente										
Construcción de clínicas de salud	1	0	6	0	7	1	2	0	9	26
Capacitación de personal de la salud	1053	0	719	0	425	1003	20	80	1085	4385
Niños que recibieron complementos alimenticios	10195	0	2240	2400	0	0	0	0	251402	266237
Niños que recibieron ayuda financiera para tratamiento dental	984	0	396	0	305	0	581	0	17	2283
Aprender										
Maestros capacitados durante una semana	0	0	367	0	970	115	565	0	303	2320
Cuadernos escolares adquiridos/donados	667	0	0	41200	0	69106	0	150	0	111123
Niños que recibieron becas	12321	0	1598	0	154	0	0	0	2014	16087
Libros escolares donados/adquiridos	0	0	45650	9600	1182	8769	7285	150	58387	131023
Pupitres escolares construidos/donados	3200	0	3689	260	1564	1725	1794	0	4109	16341
Salones de clase construidos permanentemente	44	0	50	8	93	31	45	0	82	353
Salones reparados	0	0	34	0	0	14	0	0	33	81
Adultos que fueron alfabetizados	1160	0	9000	568	3617	0	0	0	350	14695
Hábitat										
Letrinas o retretes excavados/fabricados	50	0	2403	0	57	162	23	96	4311	7102
Casas con servicio de drenaje	143	0	0	0	0	0	0	0	0	143
Pozos mejorados	0	0	15	0	7	13	0	0	159	194
Nuevos pozos (excavados)	0	0	8	93	14	0	27	0	220	362
Nuevas redes de agua potable	0	0	28	0	1	0	0	0	0	29
Redes de agua potable reparadas	0	0	392	0	2	0	0	0	31	425
Casas remozadas	265	0	520	0	0	0	1	0	2	788
Nuevas casas construidas	225	0	596	0	0	2	6	0	313	1142
Centros comunitarios construidos o mejorados	2	0	2	0	3	0	3	0	2	12
Entrenamiento de líderes comunitarios >1 día	2214	95	3522	232	200	3575	814	20	2693	13365
Kilómetros de caminos reparados	1.2	0	26	0	0	0	0	0	53.4	80.6
Puentes construidos	0	0	4	2	11	0	0	0	1	18
Familias beneficiadas por el control de la erosión	0	0	1092	0	1500	0	0	0	18405	20997
Casas que ahora cuentan con servicios de electricidad	448	0	2	0	0	0	0	0	44	494

3B

Continua al reverso de la hoja

La tabla de la página anterior forma parte de un informe publicado por PLAN Internacional, una organización internacional de asistencia. En ésta se muestra alguna información sobre el trabajo de PLAN en una de sus regiones de operación (África Oriental y del Sur). Consulta la tabla para responder las siguientes preguntas. **Circula** la respuesta correcta.

5.- De acuerdo con la tabla, ¿En qué país se realizo mayor numero de actividades de PLAN Internacional?
   A. Zambia.
   B. Malawi.
   C. Kenia.
   D. Tanzania.

6.- ¿Qué dice la tabla sobre el nivel de actividad de PLAN Internacional en Etiopia en 1996, en comparación con su actividad en otros países de la región?
   A. El nivel de actividad fue comparativamente alto en Etiopia.
   B. El nivel de actividad fue comparativamente bajo en Etiopia.
   C. Fue más o menos el mismo que en otros países de la región.
   D. Fue comparativamente alto en la categoría de hábitat y bajo en otras categorías.

7.- En 1996 Etiopia fue uno de los países más pobres en el mundo. Tomando en cuenta este hecho y la información de la tabla, ¿Qué crees que podría explicar el nivel de las actividades de PLAN Internacional realizadas en Etiopia, en comparación con sus actividades en otros países?

_____

_____

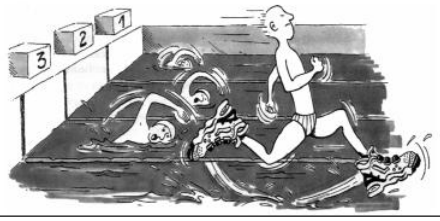
_____

8.- A continuación encontrarás algunas otras actividades de asistencia. Si estas actividades se añadieran a la tabla, ¿a qué categoría pertenecería cada una? Indica tu respuesta **cruzando** el recuadro correcto, a continuación de cada una de las actividades.

	Crecer saludablemente	Aprender	Hábitat
Teléfonos públicos instalados	☐	☐	☐
Niños vacunados	☐	☐	☐
Madres que recibieron información sobre la alimentación de infantes	☐	☐	☐
Plantas solares de generación de energía eléctrica construidas	☐	☐	☐

4B

SIENTASE BIEN EN SUS ZAPATOS DEPORTIVOS



Durante 14 años el Centro Médico Deportivo de Lyon (Francia) ha estado estudiando las lesiones en deportistas jóvenes y profesionales. El estudio ha establecido que el mejor método es la prevención... y los buenos zapatos.

El 18% de los jugadores entre los 8 y los 12 años tienen lesiones en los talones. El cartílago del tobillo de un jugador de fútbol, no responde bien a impactos y el 25% de los profesionales han descubierto que es un punto especialmente débil. El cartílago de la articulación de la rodilla, puede ser irremediablemente dañado y si no se toman los cuidados indicados desde niño (10-12 años), puede generar osteoartritis prematura. Las caderas tampoco están exentas de daño, cuando los jugadores están cansados, corren el riesgo de fracturas como resultado de caídas o colisiones.
El estudio demuestra que los jugadores de fútbol que han jugado por más de diez años, tienen sobrecrecimientos óseos en la tibia o en el talón. Esto se conoce como "el pie de futbolero", una deformidad causada por zapatos con suelas y partes del tobillo demasiado flexibles.

Proteger, dar soporte, estabilizar, amortiguar

Si un zapato es rígido, restringe el movimiento. Si es demasiado flexible, incrementa el riesgo de lesiones y torceduras. Un buen zapato deportivo debe considerar cuatro criterios: Primero, debe proporcionar *protección externa*: resistencia a los golpes de la pelota o de otro jugador, debe lidiar con irregularidades del suelo y mantener al pie caliente y seco aún si hace una temperatura helada y llueve. Debe dar *soporte al pie*, y en particular a la articulación del tobillo, para evitar torceduras, inflamaciones y otros problemas, que pueden afectar la rodilla.

Además debe proveer al jugador de una buena *estabilidad*, no se resbala en la tierra mojada o patina en una superficie seca.
Finalmente debe *amortiguar los impactos,* especialmente aquellos sufridos por los jugadores de voleibol y basquetbol quienes están constantemente saltando.

Pies secos

Para prevenir molestias pequeñas pero dolorosas como ampollas o pie de atleta (infecciones de hongos), el zapato debe permitir la evaporación de la transpiración y debe evitar que la humedad exterior lo penetre. El material ideal es la piel, que puede ser resistente al agua para prevenir que el zapato se humedezca cuando llueve.

Continua al reverso de la hoja

5B

Consulta la información para contestar las preguntas. Circula la respuesta correcta.

9.- ¿Qué es lo que pretende demostrar el autor en este texto?

A. Que la calidad de los zapatos tenis para muchos deportes ha mejorado notablemente.
B. Que es mejor no jugar futbol si tienes menos de doce años de edad.
C. Que las personas jóvenes tienen cada vez mas lesiones dada su mala condición física.
D. Que es muy importante para los jóvenes deportistas usar buen calzado deportivo.

10.- De acuerdo con el artículo, ¿Por qué los zapatos deportivos no deben ser demasiado rígidos?

_____

11.- Mira la siguiente oración ubicada casi al final del artículo (se presenta en dos partes) :

"Para prevenir molestias pequeñas pero dolorosas tales como ampollas o rajaduras de la piel o pie de atleta (infecciones de hongos)…" **(Primera parte)**

"…el zapato debe permitir la evaporación de la transpiración y debe evitar que la humedad exterior lo penetre." **(Segunda parte)**

¿Cuál es la relación entre la primera y segunda parte de la oración?

La segunda parte:

A. Contradice a la primera parte.
B. Confirma la primera parte.
C. Ilustra el problema descrito en la primera parte.
D. Da la solución al problema descrito en la primera parte.

Delfines

Características físicas

Por Atlántida Cancún

Los delfines nariz de botella en Florida miden de 2.5 a 2.7 metros y pesan entre 190 y 260 kilogramos. Las diferencias en el tamaño están relacionadas con los diferentes hábitats. Los delfines en el Pacífico miden hasta 3.7 metros y pesan 454 kilogramos; en el mediterráneo, crecen hasta 3.7 metros o más. Los machos son más largos que las hembras y más pesados.

Un delfín nariz de botella tiene un cuerpo liso, aerodinámico, fusiforme, adaptado para la vida en el ambiente acuático. La coloración es de un gris verdoso a un gris oscuro en la parte trasera, decolorándose a blanco en el vientre, la mandíbula inferior, y las regiones anales.

Las aletas pectorales tienen los elementos esqueléticos principales de las extremidades delanteras de los mamíferos terrestres, pero son más cortas y se han modificado. Los delfines las utilizan principalmente para dirigir y frenar.
La circulación de la sangre en las aletas se ajusta para ayudar a mantener la temperatura del cuerpo. Las venas en las aletas rodean a las arterias. Así, el calor de la sangre que viaja a través de las arterias se transfiere a la sangre venosa. Este intercambio ayuda a los delfines a conservar el calor del cuerpo.
Un delfín nariz de botella tiene un rostro bien definido. Puede parecer que siempre están sonriendo, no obstante, su cara está fijada en esa expresión incluso cuando sienten dolor o peligro.

6B

Los dientes, cónicos enclavijados, están diseñados para agarrar el alimento. El número de dientes varía entre individuos. La mayoría tiene de 20 a 25 dientes en cada lado de la quijada superior y de 18 a 24 en la quijada inferior, un total de 76 a 98 dientes.

Los ojos están a los lados de la cabeza, cerca de las esquinas de la boca. Las glándulas en las esquinas internas de las cuencas del ojo secretan una sustancia aceitosa que lubrica los ojos, quita los desechos, y ayuda al ojo de manera aerodinámica mientras nada. Esta lágrima puede actuar como película protectora de los ojos contra organismos contagiosos.
Los oídos, situados al lado de los ojos, son pequeñas y discretas aberturas, sin aletillas externas.

El respiradero, situado en la superficie dorsal de la cabeza, está cubierto por una aleta muscular que proporciona un sello hermético. Un delfín respira a través de su respiradero mediante la contracción de la aleta muscular.

Consulta la información para contestar las preguntas. Circula la respuesta correcta.

12.- ¿Cuál es el propósito de que las arterias estén rodeadas por las venas?
   A. Proteger las arterias en caso de toxicidad en el agua
   B. Permitir la regulación de la temperatura
   C. Las venas, al estar más superficiales, ayudan a dar la coloración a la piel
   D. Permitir que la temperatura de las venas aumente para protegerse de sus depredadores.

13.- Anota tres funciones de la substancia producida por las glándulas situadas en los ojos.

a)_____

b)_____

c)_____

14.- ¿Qué es lo que hace el delfín nariz de botella para abrir el respiradero?

_____

_____

La casa encantada

Una joven soñó una noche que caminaba por un extraño sendero campesino, que ascendía por una colina boscosa cuya cima estaba coronada por una hermosa casita blanca, rodeada de un jardín. Incapaz de ocultar su placer, llamó a la puerta de la casa, que finalmente fue abierta por un hombre muy, muy anciano, con una larga barba blanca. En el momento en que ella empezaba a hablarle, despertó. Todos los detalles de este sueño permanecieron tan grabados en su memoria, que por espacio de varios días no pudo pensar en otra cosa. Después volvió a tener el mismo sueño en tres noches sucesivas. Y siempre despertaba en el instante en que iba a empezar su conversación con el anciano.
Pocas semanas más tarde la joven se dirigía en automóvil a Litchfield, donde se realizaba una fiesta de fin de semana. De pronto tironeó la manga del conductor y le pidió que detuviera el automóvil. Allí, a la derecha del camino pavimentado, estaba el sendero campesino de su sueño.

Continua al reverso de la hoja

7B

—Espéreme un momento —suplicó, y echó a andar por el sendero, con el corazón latiéndole alocadamente. Ya no se sintió sorprendida cuando el caminito subió enroscándose hasta la cima de la boscosa colina y la dejó ante la casa cuyos menores detalles recordaba ahora con tanta precisión. El mismo anciano del sueño respondió a su impaciente llamado.
—Dígame —dijo ella—, ¿se vende esta casa?
—Sí —respondió el hombre—, pero no le aconsejo que la compre. ¡Esta casa, hija mía, está frecuentada por un fantasma!
—Un fantasma —repitió la muchacha—. Santo Dios, ¿y quién es?
—Usted —dijo el anciano y cerró la puerta.

Anónimo

15.- ¿Hacia dónde lleva el sendero en el sueño de la autora?

A. Hacia el bosque
B. Hacia Litchfield
C. Hacia un jardín
D. Hacia una casa blanca

16.- La autora ignora que está muerta. Subraya en el texto la oración que lo demuestra.

17.- ¿Por qué quiere saber la autora si la casa está en venta?

_____

_____

Aquí esta parte de la conversación entre dos personas que leyeron "La casa encantada"

18.- Usa información de la lectura para mostrar como estas personas podrían justificar sus puntos de vista.

Persona 1

_____

_____

Persona 2

_____

_____

8B

C.7. Reading comprehension test - version C

THE UNIVERSITY of York

Nombre: _____ **Grupo:** _____

NL: _____

Las Armas Científicas de la Policía

Se ha cometido un asesinato, pero el sospechoso niega todo. El asegura no conocer a la víctima. El dice jamás haberla visto, jamás la abordo, jamás la toco… la policía y el juez están convencidos de que no dice la verdad. Pero, ¿Cómo demostrarlo?

En el lugar del crimen, los detectives han recolectado todos los indicios posibles e imaginables de evidencia: pedazos de tejido, cabellos, huellas digitales, colillas de cigarro... Algunos cabellos encontrados sobre las vestimentas de la víctima son pelirrojos. Y extrañamente se parecen a los del sospechoso. Si pudiésemos demostrar que esos cabellos son realmente suyos, sería una prueba de que él conocía a la víctima.

Cada persona es única

Los especialistas se pusieron manos a la obra. Examinaron algunas células de la raíz de los cabellos y algunas células sanguíneas del sospechoso. En el núcleo de cada una de las células de nuestros cuerpos, encontramos ADN.

¿De qué se trata esto? El ADN es como una cadena formada de dos rangos de perlas entrelazados.

Imagina que esas perlas son de cuatro colores diferentes y que miles de ellas (las cuales estructuran un gen) están alineadas de acuerdo a un orden preciso. En cada individuo, este orden es exactamente el mismo en todas las células del cuerpo, en las de la raíz del pelo, las del dedo gordo del pie, las del hígado, las del estómago o de la sangre.

Sin embargo, entre una persona y otra, el orden de las perlas varía. Visto el número de perlas que están alineadas de este modo, existe poca probabilidad de que dos personas posean el mismo ADN, a excepción de que sean auténticos gemelos. Único para cada individuo, el ADN es, de esta manera, una especie de identificación genética.

Los expertos en genética son capaces de comparar la tarjeta de identificación genética del sospechoso (determinada gracias a su sangre) y aquella de la persona del cabello pelirrojo. Si se tratara de la misma tarjeta genética, sabríamos que el sospechoso efectivamente abordó a la víctima que él dice jamás haber visto.

En los casos de abusos sexuales, homicidios, asaltos y otro tipo de delitos, la policía hace cada vez más análisis genéticos. ¿Por qué? Para tratar de encontrar pruebas que establezcan un contacto entre dos personas, entre dos objetos o entre una persona y un objeto. Hacer la prueba de dicho contacto es frecuentemente útil para la investigación. Pero no necesariamente proporciona la prueba de un crimen. Se trata solamente de un elemento de evidencia entre muchos otros.

Anne Versailles

Continua al reverso de la hoja

1C

Estamos formados por billones de células Cada ser viviente se compone de múltiples células. Una célula es infinitamente pequeña. Decimos que es microscópica porque solamente la podemos ver con un microscopio que maximiza su tamaño muchas veces. Cada célula está formada por una envoltura y un núcleo, en el cual se encuentra el ADN. ***¿Qué es la genética?*** El ADN se compone de varios genes. El conjunto de genes, estando formado cada uno de ellos por miles de perlas. Todos estos genes juntos forman la tarjeta de identificación genética de una persona. ***¿Cómo se revela la tarjeta de identificación genética?*** Los expertos en genética toman algunas	células que se encuentran en la raíz del cabello de la víctima o en la saliva restante en las colillas de los cigarros. La sumergen en una solución que destruye todo lo que rodea el ADN de estas células. Hacen lo mismo con las células de la sangre del sospechoso. El ADN se prepara especialmente para el análisis. Después de eso, se pone sobre una gelatina especial. Enseguida se le conecta una corriente de electricidad a la gelatina. Después de algunas horas, se producen unas bandas (parecidas a los códigos de barras que encontramos en los productos que compramos), visibles solamente bajo una lámpara especial. Comparamos entonces el código de barras del ADN del sospechoso y aquél de los cabellos encontrados sobre la víctima.

Consulta el artículo 'Las Armas Científicas de la Policía' para responder las siguientes preguntas. Circula la respuesta correcta.

1.- Al conocer la tarjeta de identificación genética de un individuo se pretende:
   A. Identificar los genes de una persona
   B. Obtener el código de barras de las personas
   C. Estudiar las uniones y colores de las cadenas de perlas
   D. Comparar su información con la de otro individuo para ver si ha habido contacto entre ellos.

2.- ¿Cuál es el objetivo principal del autor?
   A. advertir.
   B. divertir.
   C. informar.
   D. convencer

3.- Menciona dos ventajas que nos da el conocer la tarjeta de identificación genética de los individuos.

a)_____

b)_____

4.- El final de la introducción (en la primera sección sombreada) dice: "Pero, ¿cómo demostrarlo?" De acuerdo al pasaje, los investigadores tratan de encontrar una respuesta a esta pregunta. Subraya en el texto, la oración que explica que es lo que hacen los investigadores.

2C

Consulta la tabla y resuelve las preguntas de la siguiente página.

Resultados del Programa Iberoamérica Crece durante el Año Fiscal de 2002

Región Centro América

	Costa Rica	Panamá	Ecuador	Belice	Honduras	República Dominicana	Guatemala	Nicaragua	El Salvador	Total
Crecer saludablemente										
Construcción de clínicas de salud	7	1	0	0	2	1	3	2	1	17
Capacitación de personal de la salud	1000	890	0	1100	3250	2370	1430	750	800	11590
Niños que recibieron complementos alimenticios	2500	1900	0	3000	859	427	890	420	350	10346
Niños que recibieron ayuda financiera para tratamiento dental	100	70	0	78	85	62	58	50	0	503
Aprender										
Maestros capacitados durante una semana	1120	860	0	0	2000	975	430	250	380	6015
Cuadernos escolares adquiridos/donados	300	150	0	0	225	180	0	285	0	1140
Niños que recibieron becas	20	15	0	3	35	0	15	12	9	109
Libros escolares donados/adquiridos	890	7000	0	130	170	260	238	154	195	9037
Pupitres escolares construidos/donados	200	120	0	0	450	348	360	130	0	1608
Salones de clase construidos permanentemente	40	10	0	5	8	10	8	11	2	94
Salones reparados	0	30	0	12	45	32	50	10	0	179
Adultos que fueron alfabetizados	300	120	0	150	0	40	95	35	25	765
Hábitat										
Letrinas o retretes excavados/fabricados	40	30	0	30	13	25	48	67	80	333
Casas conectadas a nuevo servicio de drenaje	450	280	0	84	480	690	230	176	82	2472
Pozos mejorados	13	1	0	0	0	8	18	23	5	68
Nuevos pozos (excavados)	6	1	0	2	0	5	4	8	3	29
Nuevas redes de agua potable	5	0	0	2	0	13	5	8	0	33
Redes de agua potable reparadas	14	4	0	1	0	2	2	0	1	24
Casas remozadas	215	30	0	20	0	15	35	60	10	385
Nuevas casas construidas	300	0	0	0	250	120	95	80	70	915
Centros comunitarios construidos o mejorados	25	22	0	10	0	20	14	0	8	99
Entrenamiento de líderes comunitarios >1 día o más	520	234	115	100	470	340	290	150	80	2299
Kilómetros de caminos reparados	35.6	40.8	0	32	30	0	0	34	0	172.4
Puentes construidos	4	2	0	1	0	0	0	2	0	9
Familias beneficiadas por el control de la erosión	450	380	0	0	120	0	232	0	80	1262
Casas que ahora cuentan con servicios de electricidad	800	423	0	370	265	170	230	130	120	2508

Continua al reverso de la hoja

3C

La tabla de la página anterior forma parte de un informe publicado por el programa Iberoamérica Crece, una organización internacional de asistencia. En ésta se muestra alguna información sobre el trabajo del programa en una de sus regiones de operación (América Central). Consulta la tabla para responder las siguientes preguntas. **Circula** la respuesta correcta.

5.- De acuerdo con la tabla, ¿En qué país se realizo mayor numero de actividades?
   A. El Salvador
   B. Costa Rica
   C. Honduras
   D. Belice

6.- ¿Qué dice la tabla sobre el nivel de actividad del programa en Ecuador en el 2002, en comparación con su actividad en otros países de la región?
   A. El nivel de actividad fue comparativamente alto en Ecuador.
   B. El nivel de actividad fue comparativamente bajo en Ecuador.
   C. Fue más o menos el mismo que en otros países de la región.
   D. Fue comparativamente alto en la categoría de hábitat y bajo en otras categorías.

7.- En el 2002 Ecuador fue uno de los países más pobres en Latinoamérica. Tomando en cuenta este hecho y la información de la tabla, ¿Qué crees que podría explicar el nivel de las actividades del programa realizadas en Ecuador, en comparación con sus actividades en otros países?

_____

_____

8.- A continuación encontrarás algunas otras actividades de asistencia. Si estas actividades se añadieran a la tabla, ¿a qué categoría pertenecería cada una? Indica tu respuesta **cruzando** el recuadro correcto, a continuación de cada una de las actividades.

	Crecer saludablemente	Aprender	Hábitat
Teléfonos públicos instalados	☐	☐	☐
Niños vacunados	☐	☐	☐
Madres que recibieron información sobre la alimentación de infantes	☐	☐	☐
Plantas solares de generación de energía eléctrica construidas	☐	☐	☐

EL TÉ PODRÍA MEJORAR LA MEMORIA

Un equipo de científicos de la University of Newcastle upon Tyne investiga formas alternativas de tratamiento de la enfermedad del Alzheimer. Los resultados de laboratorio indican que ciertas variedades de té inhiben la actividad de determinadas enzimas en el cerebro que están asociadas con la memoria.

4C

En sus experimentos, los científicos del Medicinal Plant Research Centre investigaron las propiedades del café, del té negro y del té verde. Los resultados sugieren que tanto el té verde como el té negro inhiben la actividad de enzimas asociadas con el desarrollo del Alzheimer. En cambio, el café no parece tener tal efecto. Ambas clases de té inhiben la actividad de la enzima acetilcolinesterasa, que descompone el mensajero químico o neurotransmisor acetilcolina. Precisamente, el Alzheimer se caracteriza por una disminución de la acetilcolina.

El té verde y el té negro también entorpecen la actividad de la enzima butilrilcolinesterasa, que ha sido descubierta en depósitos de proteínas de pacientes con Alzheimer. De hecho, el té verde obstruye la actividad de la beta-secretasa, una sustancia que participa en la producción de los depósitos de proteínas en el cerebro asociados a la enfermedad. Según los experimentos, este té prolonga su efecto inhibidor durante una semana, mientras que el té negro sólo lo hace durante un día.

Por el momento no hay cura para el Alzheimer, pero es posible hacer más lento su desarrollo. Los fármacos actualmente disponibles entorpecen la actividad de las enzimas antes mencionadas. Sin embargo, muchas de estas drogas, como el donepezil, tienen efectos secundarios, por lo que los médicos están buscando alternativas.

Los investigadores de la Newcastle University, encabezados por Ed Okello, creen que sus trabajos en relación con el té pueden ser muy útiles en este campo. Mientras tanto las personas que quieran prevenir el Alzheimer harían bien en consumir esta bebida tradicional, ya que tras el consumo del té verde no sólo conseguirán prevenir la enfermedad, sino también mantendrán una memoria fantástica, incluso a una avanzada edad.

Consulta la información para contestar las preguntas. Circula la respuesta correcta.

9.- ¿Qué es lo que pretende el autor en este texto?

A. Informar sobre el beneficio del té para la memoria.
B. Estudiar las propiedades tanto del té verde como del té negro.
C. Publicar que algunos tés entorpecen la actividad de algunas enzimas.
D. Dar a conocer la investigación de un grupo de científicos sobre el tratamiento adecuado para el Alzheimer.

10.- De acuerdo con el artículo, ¿Por qué las personas que consuman té verde y té negro previenen el Alzheimer?

_____

_____

11.- Mira la siguiente oración ubicada casi al final del artículo (se presenta en dos partes):

"Mientras tanto las personas que quieran prevenir el Alzheimer harían bien en consumir esta bebida tradicional…" **(Primera parte)**

"…ya que tras el consumo del té verde no sólo conseguirán prevenir la enfermedad, sino también mantendrán una memoria fantástica." **(Segunda parte)**

¿Cuál es la relación entre la primera y segunda parte de la oración?

Continua al reverso de la hoja

5C

La segunda parte:

A. Contradice a la primera parte.
B. Confirma la primera parte.
C. Ilustra el problema descrito en la primera parte.
D. Da la solución al problema descrito en la primera parte.

LAS NUBES

La superficie terrestre está cubierta en un 70% de agua. Cuando el Sol calienta el agua de los mares, ríos, lagos y estanques, provoca la evaporación de una gran cantidad de agua; al mismo tiempo las plantas verdes absorben la humedad del terreno y la introducen en el aire, a través de las hojas, bajo la forma de vapor acuoso. Las nubes se forman para modificar la distribución del calor solar sobre la superficie terrestre y en la atmósfera.

Formación de las nubes
El aire húmedo y caliente producido en la superficie de la atmósfera se eleva poco a poco y se va enfriando. En cierto punto de enfriamiento, el vapor acuoso se condensa en microscópicas gotitas. Las nubes están constituidas por estas diminutas gotas.
Tipos de nubes

Las nubes se clasifican según distintas características, como su forma, la altura en que se encuentran y su estado material. En cuanto a la forma se distinguen tres tipos fundamentales de nubes: estratos, cúmulos y cirros.

Estratos
Son las nubes más bajas. Están en posición horizontal y son de color blanco o gris tenue. Los estratos se ven a menudo como que cubren todo el cielo. Como son tan delgados, rara vez producen mucha lluvia o nieve. A veces, en las montañas o colinas, estas nubes se confunden con neblina.

Cúmulos
Son nubes medias. Tienen un aspecto esponjoso, semejante a bolas de algodón con contornos bien definidos. Las soleadas partes superiores de las nubes son de un blanco brillante, mientras que su base horizontal se muestra sombría y oscura. Se forman normalmente durante el día y se disuelven de noche. En general son un signo de buen tiempo.

Cirros
Son nubes altas. Tienen una apariencia de hebras delgadas y alargadas. Durante el día son blanquísimos; al atardecer, al incidir sobre ellos la luz crepuscular, presentan un color amarillento o rojizo. Se encuentran a más de 7000 metros de altura donde hay temperaturas bajo cero, de tal forma que los cirros son nubes de hielo.
Además de los tres tipos fundamentales de nubes, existen nubes que poseen características mixtas como, por ejemplo, los cirrocúmulos, los cirroestratos y los estratocúmulos. En total existen 10 tipos de nubes establecidos por la Organización Mundial Meteorológica (WMO). Sin embargo, sus límites no son siempre exactos, pues las nubes se mueven con tal libertad que los rebasan frecuentemente.

6C

Consulta la información para contestar las preguntas. Circula la respuesta correcta.

12.- ¿Cuál es el propósito de la formación de las nubes?

A. Absorber la humedad del aire
B. Modificar la distribución del calor en la atmósfera
C. Mantener fresca la superficie terrestre mediante la creación de sombras
D. Indicar el buen o mal tiempo

13.- Anota tres de las características que ayudan a clasificar a las nubes.

a)_____

b)_____

c)_____

14.- ¿Cómo se forman las nubes?

_____

_____

LA SANGRE DE MEDUSA

Termina el otoño. Las calles de la ciudad se cubren de hojas secas. Las aves emigran al sur. Al atardecer cruzan la ciudad junto al lago. Arde el sol en las ventanas de los edificios más altos. Cegados ante el resplandor muchos pájaros se estrellan contra los cristales y caen muertos. Otros quedan malheridos. A menudo la agonía termina en el basurero municipal o entre las llamas de los incineradores.

Todas las noches Jack recorre las calles en busca de aves caídas. Arroja los cadáveres en un costal, para sepultarlos en su jardín, y pone en cajitas forradas de algodón a las aves que encuentra aún con vida. Su departamento está lleno de pájaros en distintas fases de convalecencia. Algunos se entrenan para recobrar la facultad del vuelo. Otros apenas dan pasos inciertos. Jack los cura, los cuida y alimenta. En medicamentos, alpiste, vitaminas y en mantener el sitio a una temperatura adecuada, gasta cuanto obtiene como redactor.

En su casa solamente están las incubadoras y una radio utilizada para enterarse del clima. Los únicos libros son de ornitología y veterinaria. Para Jack significa una tragedia la muerte de un pájaro. Jack es dueño de un talento médico natural y la práctica le da una destreza incomparable. El suyo es el amor perfecto: no exige retribución, aplauso ni alabanza. Lo hace feliz abrir la ventana y dejar que las aves reanuden el vuelo rumbo al sur.

Hoy la temperatura ha descendido a cero. Jack sobrevuela la ciudad junto al lago. En lo alto encuentra una dicha desconocida aquí abajo. Al fin sabe qué son el júbilo y el poder de los pájaros, sentimientos tan opuestos a la angustia y la indefensión de los seres humanos. Quiere decir algunas palabras: sólo gorjeos brotan de su pico. Su amor al fin lo ha convertido en el objeto amado. Pero el sol muriente lo enceguece. Jack va a estrellarse sobre el pavimento. Sólo por las plumas será posible reconocer su cadáver.

José Emilio Pacheco

7C

Continua al reverso de la hoja

15.- ¿En qué época del año se desarrolla la historia?

A. primavera
B. verano
C. otoño
D. invierno

16.- ¿Qué es lo que le sucede a Jack al final de la historia? Subraya en el texto la oración que lo demuestra.

17.- ¿Por qué se dedica Jack a recoger las aves?

_____

_____

Aquí esta parte de la conversación entre dos personas que leyeron "La Medusa"

18.- Usa información de la lectura para mostrar como estas personas podrían justificar sus puntos de vista.

Persona 1

_____

_____

Persona 2

_____

_____

8C

C.8. Metacognitive assessment instrument (MAI)

THE UNIVERSITY of York

Nombre: _____ **Grupo:** _____

NL: _____

¿Y tú qué haces cuando lees?

A continuación hay una lista de frases acerca de lo que las personas hacen cuando leen textos académicos o materiales relacionados con la escuela como *libros de texto*, o *libros de la biblioteca*. Delante de cada frase hay cuatro casillas numeradas que significan lo siguiente:

1 – "Yo **nunca** o **casi nunca** hago esto."
2 – "Yo hago esto sólo **de vez en cuando**."
3 – "Yo **normalmente** hago esto." (como el 50% del tiempo)
4 – "Yo **siempre** o **casi siempre** hago esto."

Lee cada frase y ***tacha*** el número que expresa lo que tú haces. NO hay respuestas correctas o incorrectas, cada persona trabaja de forma distinta.

Continua al reverso de la hoja

1

		nunca o **casi nunca**	**de vez en cuando**	**normalmente**	**siempre** o **casi siempre**
1.	Antes de comenzar a leer pienso en la razón por la que voy a leer para ayudarme durante la lectura.	1	2	3	4
2.	Reviso el texto de forma rápida primero para darme cuenta de su organización, extensión y otras características estilísticas.	1	2	3	4
3.	Pienso acerca de lo que ya sé para ayudarme a entender mejor lo que leo.	1	2	3	4
4.	Mientras leo me hago preguntas como si mi profesor me fuera a hacer un examen.	1	2	3	4
5.	Cuando leo un texto o una frase y no lo entiendo vuelvo a leerlo para darme cuenta de **por qué** no lo entiendo.	1	2	3	4
6.	Cuando leo un texto, lo leo varias veces para aumentar mi comprensión.	1	2	3	4
7.	Cuando me encuentro durante la lectura con **una frase** cuyo significado no entiendo, leo las anteriores así como las posteriores para ayudarme a comprender el significado de la frase que no entendí.	1	2	3	4
8.	Trato de volver atrás y encontrar nuevamente la idea cuando pierdo la concentración.	1	2	3	4
9.	Ajusto mi velocidad de lectura de acuerdo a lo que estoy leyendo y al por qué lo estoy leyendo (ej. si es un texto difícil, lo leo lentamente tratando de comprenderlo).	1	2	3	4
10.	Uso las pistas (ej. palabras claves) en el texto para ayudarme a comprender mejor lo que leo.	1	2	3	4
11.	Mientras leo, trato de descubrir las partes de que se compone el texto (introducción, nudo, desenlace, etc.)	1	2	3	4
12.	Utilizo marcadores o subrayo información en el texto para ayudarme a recordar ideas y conceptos importantes de un texto.	1	2	3	4
13.	Me auxilio de estilos tipográficos (letras negritas o cursivas) para ayudarme a identificar información importante.	1	2	3	4
14.	Cuando un texto es difícil de leer, pregunto a otras personas (profesores o amigos) para ayudarme a comprenderlo.	1	2	3	4
15.	Organizo la información que he leído para ayudarme a comprender mejor.	1	2	3	4
16.	Cuando termino de leer un texto me aseguro de que puedo decir (a alguien o a mi mismo) con mis palabras de lo que se trata el texto.	1	2	3	4
17.	Pienso si el contenido del texto cumple con mi propósito de leer.	1	2	3	4
18.	Analizo críticamente y evalúo la información presentada en el texto.	1	2	3	4
19.	Reviso si mis predicciones acerca del texto fueron correctas o incorrectas.	1	2	3	4

2

		nunca o casi nunca	de vez en cuando	normalmente	siempre o casi siempre
20.	Cuando leo trato de adivinar de qué se trata el material.	1	2	3	4
21.	Trato de visualizar la información en mi mente para ayudarme a recordar lo que leo.	1	2	3	4
22.	Cuando leo tengo en mente las estrategias de lectura que pueden ayudarme a comprender mejor.	1	2	3	4
23.	Antes de empezar a leer reviso el texto, veo el título y las ilustraciones, para ver de qué se trata.	1	2	3	4
24.	Me hago preguntas que me gustaría que el texto contestara.	1	2	3	4
25.	Vuelvo a leer el texto cuando tengo que elaborar conclusiones sobre lo que he leído.	1	2	3	4
26.	Voy hacia adelante y hacia atrás en el texto para encontrar las relaciones entre las ideas del texto.	1	2	3	4
27.	Si hay una **palabra** que no entiendo, releo las palabras anteriores y leo las posteriores para ayudarme a entender.	1	2	3	4
28.	Yo decido que leer con detenimiento y que ignorar.	1	2	3	4
29.	Uso las tablas, figuras y dibujos en el texto para facilitar mi comprensión.	1	2	3	4
30.	Trato de entender el significado de palabras o frases desconocidas mediante las pistas que encuentro en el texto.	1	2	3	4
31.	Cuando encuentro un párrafo difícil de entender continuo leyendo para ver si se aclara el significado más adelante.	1	2	3	4
32.	Me detengo de vez en cuando para pensar acerca de lo que estoy leyendo.	1	2	3	4
33.	Uso materiales de referencia como diccionarios, enciclopedias o la computadora para ayudarme a comprender lo que leo.	1	2	3	4
34.	Tomo notas o escribo en los márgenes mientras leo para ayudarme a entender lo que leo.	1	2	3	4
35.	Resumo un texto para seleccionar la información importante del mismo.	1	2	3	4
36.	Pongo las ideas con mis propias palabras (parafraseo) para entender mejor lo que leo.	1	2	3	4
37.	Comento lo que leo con otras personas.	1	2	3	4
38.	Cuando encuentro alguna oración que no concuerda con el texto o que parece destacar, reviso lo que ya he leído.	1	2	3	4
39.	Mientras leo me detengo a pensar lo que va a pasar más adelante o a que conclusiones llegará el/la autor/a.	1	2	3	4
40.	Cuando leo busco las partes más importantes del texto.	1	2	3	4

3

Continua al reverso de la hoja

		nunca o casi nunca	de vez en cuando	normalmente	siempre o casi siempre
41.	Sé en qué momento (***cuando***) utilizar las estrategias de lectura.	1	2	3	4
42.	Al terminar de leer pienso en lo que aprendí del texto.	1	2	3	4
43.	Trato de relacionar la información del texto con mis otras materias de la escuela.	1	2	3	4
44.	Me imagino los eventos del texto para facilitar su comprensión.	1	2	3	4
45.	Leo los textos una sola vez.	1	2	3	4
46.	Reviso si entendí lo que he leído y me pregunto qué conceptos todavía no he entendido.	1	2	3	4
47.	Cuando leo, resalto los detalles de la lectura.	1	2	3	4
48.	Cuando leo pienso en como la información del texto tiene relación con la vida real.	1	2	3	4

4

C.9. Text used in verbal report (post-test)

La frigidez del mármol

Alguna vez he sentido escalofrió al ver las estatuas del parque; en especial las de mármol. Si al caminar por un espacio al aire libre uno se las encuentra, puede que un frío intenso nos invada por dentro al pensar en tan bellas redondeces expuestas a la inclemencia del clima. Mientras nosotros disfrutamos de prendas abrigadoras, ellas permanecen impávidas con su inmutable actitud. Sin duda nuestro estado original, la desnudez, fue un acertado ivento de la madre naturaleza; en tal condición no requeríamos los innumerables catálogos y asesores que conforman hoy parte de la historia de la moda la cual fue creada desde hace varios miles de arcos comenzando con las pieles del hombre prehistórico. Y peludos, al estilo de los orangutanes, vagábamos por el paraíso tan quitados de la pena; así, sin pena, pero habría de llegar el buen Dios a ponernos "de patitas en el mundo" y colgarnos al rostro la vergenza de nuestra desnudez.

Los estudiosos explican a su entender la pérdida del bello corporal en el homo sapiens: Desmond Morris en El Mono Desnudo la califica como benéfica pues según argumenta, al quedarnos pelambre sin dejamos en ella los bichos que nos invadían viviendo en las cavernas. Otro antropólogo, José H. Reichholf en su obra La Aparición del Hombre sostiene algo de mayor trascendencia pues afirma que sólo desnudos podemos sudar suficientemente para refrigerar nuestro cuerpo, es decir, para regular la temperatura corporal ya que nuestra especie se distinguió del resto de los primates por la condición de bípedos corredores. Además, según este estudioso, logramos desarrollar nuestra actividad cerebral debido al fósforo que contiene la carne que consumimos desde hace milenios. De manera que nuestra costumbre carroñera a similitud de los buitres nos obligó a competir con ellos para ganar una presa y esto solo era posible corriendo... y sudando pues solamente así, sudando nuestro organismo es capaz de correr más de una obra. Pues la realidad es que vamos por el mundo sin tener ideas claras de cómo manejar nuestra desnudez pero es bien claro que hechos tan importantes como nacer; tan trascendentes como amar o tan divertidos como sumergirnos en el agua para nadar, los hacemos desnudos; quizá conscientes de que la desnudez es un vinculo ancestral. Tal vez por ello nos impactamos cuando en una playa o en pleno Paseo de la Reforma hombres o mujeres desnudos abofetean nuestras pudorosas costumbres hartas de tantos trapos.

Pues ya sea andar "bichi", "chirundo" o "en pelotas" como dicen le en Argentina, nuestra piel requiere de baños de sol, vitamina E y una dosis de caricias para no parecer una frígida estatua de mármol olvidada en el medio del parque.

C.10. Text used in verbal report (post-test 2)

Mitos, leyendas, cuentos, fábulas, apólogos y parábolas

Las mariposas que hoy vemos ingrávidas, que se pueden posar en las flores, en la superficie de las aguas y hasta en las trémulas ramas del aire, no son otra cosa que la fracasada imagen de lo que el murciélago fue en otro tiempo: el ave más bella de la creación. Pero no siempre fue así: cuando la luz y la sombra echaron a andar, era como ahora lo conocemos y se llamaba biguidibela: biguidi, mariposa y bela, carne: mariposa en carne, es decir, desnuda. La más fea y más desventurada de todas las criaturas era entonces el murciélago. Y un día, acosado por el frío, subió al cielo y dijo a Dios:

- Me muero de frío. Necesito de plumas.

Y como Dios, aunque no cesa de trabajar tampoco vuelve las manos a tareas ya cumplidas, tenía no ya ninguna pluma. Así fue que le dijo que volviera a la Tierra y suplicara en su nombre una pluma a todas las naves. Porque Dios da siempre más de lo que se le pide. Y el murciélago, vuelto a la Tierra, recurrió a los pájaros de más vistoso plumaje. La pluma verde del cuello de los loros, la azul de la paloma azul, la blanca de la paloma blanca, la tornasol de la chuparrosa, su más próxima imagen actual: todas las tuvo el murciélago. Y orgulloso volaba sobre las sienes de la maana, y las otras aves, refrenando el vuelo, se detenían para admirarlo. Y había una emoción nueva, plástica, sobre la Tierra. A la caída de la tade, volando con el viento del poniente, coloreaba el horizonte. Y una vez, viniendo de más allá de las nubes, creó el arco iris, como un eco de su vuelo. Sentado en las ramas de los árboles abría alternativamente las alas, sacudiéndolas en un temblor que alegraba el aire. Todas las aves comenzaron a sentir envidia de él y el odio se volvió unánime, como un día lo fue la admiración.

Otro día subió al cielo una pájaros de parvada, el colibrí iba adelante. Dios oyó su queja. El murciélago se burlaba de ellos; además, con una pluma menos padecían frío. Y ellos mismos trajeron el mensaje celestial en el que se le llamaba al murciélago. Cuando estuvo en la casa de allá arriba, Dios le hizo repetir los ademanes que habían ofendido a sus compaeros; y agitando las alas se quedo otra vez desnudo. Se dice que todo un día llovieron plumas del cielo.

Y desde entonces sólo vuela en los atardeceres en rápidos giros, cazando uas imaginarias. Y no se detiene, para que nadie advierta su fealdad.

C.11. Background questionnaire

THE UNIVERSITY of York

CUESTIONARIO PARA ALUMNOS

El siguiente cuestionario es para conocer un poco más acerca de ti y de tu educación. Contesta lo que se te pide o marca las casillas ☒ que reflejen tu forma de pensar. Si tienes duda, por favor pregúntale a la persona que está aplicando el cuestionario. Gracias por tu participación.

Nombre: _____

Grupo: _____ **No. de lista:** _____

1.- ¿Cuál es tu fecha de nacimiento? *(Por favor escribe día, mes y año)*

_____ _____ 19_____
día mes año

2.- ¿Eres hombre o mujer?
☐ hombre ☐ mujer

3.- ¿Cursaste preprimaria o preescolar?
☐ No
☐ Si, por un año o menos
☐ Si, por más de un año

4.- ¿Qué hace tu madre en su trabajo? *(Ej. Enseña a alumnos de primaria, ayuda a preparar comida en un restaurante, vende cosméticos, etc.)*

_____

5.- ¿Qué hace tu padre en su trabajo? *(Ej. Construye casas, maneja un taxi, diseña muebles, etc.)*

_____

6.- ¿Cuál de los siguientes objetos se puede encontrar en tu casa? *(Elige una opción en cada línea)*

SI	NO	
☐	☐	a) un escritorio para estudiar
☐	☐	b) un cuarto para ti solo
☐	☐	c) un lugar tranquilo para estudiar
☐	☐	d) una computadora que puedes usar para trabajos relacionados con la escuela
☐	☐	e) programas de computadora (software) con temas relacionados con la escuela
☐	☐	f) conexión a internet
☐	☐	g) libros de literatura clásica
☐	☐	h) libros de poesía
☐	☐	i) piezas de arte (ej. pinturas, esculturas)
☐	☐	j) libros para apoyarte con tus tareas en casa
☐	☐	k) enciclopedias
☐	☐	l) diccionarios
☐	☐	m) DVD
☐	☐	n) Nintendo/WIFI

Continua al reverso de la hoja

1

7.- ¿Cuántos libros hay en tu casa? No incluyas revistas, periódicos ni tus libros de la escuela.

- ☐ 0 – 10 libros
- ☐ 11 – 25 libros
- ☐ 26 – 100 libros
- ☐ Más de 100 libros

8.- ¿Cuánto tiempo pasas normalmente leyendo por placer? *(Elige sólo una opción)*

- ☐ No leo por placer
- ☐ 30 minutos o menos al día
- ☐ Entre 30 y 60 minutos cada día
- ☐ 1 a 2 horas al día
- ☐ Más de 2 horas al día

9.- ¿Qué tan de acuerdo o en desacuerdo estás con las siguientes oraciones acerca de la lectura?
(Elige sólo una opción en cada línea)

		Totalmente en desacuerdo	En desacuerdo	De acuerdo	Totalmente de acuerdo
a)	Leo sólo si lo tengo que hacer.	☐	☐	☐	☐
b)	Leer es uno de mis pasatiempos favoritos.	☐	☐	☐	☐
c)	Me gusta hablar de libros con otras personas.	☐	☐	☐	☐
d)	Se me hace difícil terminar de leer libros.	☐	☐	☐	☐
e)	Me siento feliz si recibo un libro como regalo.	☐	☐	☐	☐
f)	Para mí, leer es una pérdida de tiempo.	☐	☐	☐	☐
g)	Me gusta ir a las librerías o a la biblioteca.	☐	☐	☐	☐
h)	Leo sólo para obtener información que necesito.	☐	☐	☐	☐
i)	No puedo sentarme quieto y leer por más de 10 minutos.	☐	☐	☐	☐
j)	Me gusta expresar mis opiniones sobre libros que he leído.	☐	☐	☐	☐
k)	Me gusta intercambiar libros con mis amigos.	☐	☐	☐	☐
l)	Me considero un buen lector.	☐	☐	☐	☐

2

10.-¿Qué tan seguido lees los siguientes tipos de materiales por gusto? *(Elige sólo una opción en cada línea)*

		Nunca o casi nunca	Algunas veces al año	Una vez al mes	Varias veces al mes	Varias veces a la semana
a)	Revistas	☐	☐	☐	☐	☐
b)	Comics	☐	☐	☐	☐	☐
c)	Libros de ficción (cuentos, novelas, aventuras)	☐	☐	☐	☐	☐
d)	Libros de no ficción (narraciones reales, textos informativos)	☐	☐	☐	☐	☐
e)	Periódicos	☐	☐	☐	☐	☐

11.-En tus clases ¿qué tan seguido ocurre lo siguiente? *(Elige sólo una opción en cada línea)*

		Nunca o casi nunca	En algunas clases	En la mayoría de las clases	En todas las clases
a)	El profesor pregunta a los alumnos que expliquen las ideas de un texto.	☐	☐	☐	☐
b)	El profesor hace preguntas que ayudan a los alumnos a comprender mejor un texto.	☐	☐	☐	☐
c)	El profesor da a los alumnos el tiempo suficiente para pensar sus respuestas.	☐	☐	☐	☐
d)	El profesor recomienda un libro o autor para leer.	☐	☐	☐	☐
e)	El profesor alienta a los alumnos a expresar sus opiniones sobre un texto.	☐	☐	☐	☐
f)	El profesor ayuda a los alumnos a relacionar las historias que leen con sus experiencias.	☐	☐	☐	☐
g)	El profesor enseña a los alumnos cómo la información de un texto los ayuda a aprender más sobre lo que ya saben.	☐	☐	☐	☐
h)	El profesor pide a los alumnos que mientras leen traten de adivinar lo que pasará más adelante en el texto.	☐	☐	☐	☐
i)	El profesor explica a los alumnos cómo las tablas, dibujos y palabras claves pueden ayudarlos a comprender el significado de palabras o frases difíciles de entender.	☐	☐	☐	☐
j)	El profesor pide a los alumnos que realicen un resumen de lo que leyeron.	☐	☐	☐	☐

Continua al reverso de la hoja

12.-¿Qué tan seguido visitas una biblioteca para realizar las siguientes actividades? *(Puede ser la de tu escuela o una cerca de tu casa)*

		Nunca	Algunas veces al año	Una vez al mes	Varias veces al mes	Varias veces a la semana
a)	Sacar libros para leer por placer.	☐	☐	☐	☐	☐
b)	Sacar libros para realizar trabajos escolares.	☐	☐	☐	☐	☐
c)	Hacer la tarea, buscar información o hacer proyectos escolares.	☐	☐	☐	☐	☐
d)	Leer revistas y periódicos.	☐	☐	☐	☐	☐
e)	Leer libros por diversión.	☐	☐	☐	☐	☐
f)	Aprender sobre actividades no relacionadas con materias de la escuela como deportes, pasatiempos, gente, música.	☐	☐	☐	☐	☐
g)	Usar el internet.	☐	☐	☐	☐	☐

13.-¿Hay biblioteca en tu escuela?

☐ Si
☐ No

14.- Al terminar la secundaria planeo …

☐ a) seguir estudiando (ej. preparatoria, CCH, CONALEP, vocacional, etc.)
☐ b) buscar trabajo
☐ c) todavía no tengo planes
☐ d) otro (anótalo) _____

4

C.12. Examples of actvities carried out with the strategy and control groups

Text	Activities in strategy group	Activities in control group
No todos vuelven a la escuela	Teacher introduces first steps of the hand:	Read story
	- Why to read (¿Por que leer?)	Mention three interest facts about the story
	- Quick scan (Reviso rápido)	Perform activities
	- What is it about? (¿De qué se trata?)	Discuss what they would do if they were Abraham (character of the story)?
	Researcher explains why to look at titles, subtitles, drawings, charts.	Discussion about their objectives for this academic year. Challenges and how to achieve their goals.
El Balcón	Practise strategies with promp Read forwards and backwards to guess meaning of words(semi-controlled activity)	Read and answer questions Look for words in dictionary Give an ending to the story
El vuelo de Lindbergh	Practise strategies with promp	Name the inventions
	Main ideas/ paraphrase (controlled activity)	Read and answer questions Design their own invention

C.13. First coding of verbal report

	Strategy (as labelled on prompt)	Description
1.	Why do I have to read?	Think about the main reason for reading (e.g. setting goals)
2.	Quick review	Students scan the text (e.g. titles, pictures, bold, etc.) to help them create an initial inference of the topic.
3.	What is it about?	Explain what the text is about based on the information gathered from the first strategy.
4.	Review and Imagine	Read half of the text and stop to try to create a personal summary of what has happened at the beginning of the text (visualise). Think about what comes next in the text (inferences). Finish reading the text and check if inferences were right.
5.	Problems (Supervise)	Students check their understanding of the text and identify problems that might be causing problems.
6.	Rereading - Read backwards, forwards	Students might not understand a word or a phrase so they should reread it, read it at a different speed or read before and after the phrase/word to solve their problem.
7.	Use of context	If some words or phrases are unknown, students could use the information in the text to try and infer the meaning.
8.	Summarise - Main ideas - Paraphrase	Students try to explain with their own words the main ideas of the text in order to make it easier to understand.
9.	Monitor	Students monitor their reading. The use of this strategy is evident when students realise that what they are reading does not make sense when compared against either their previous knowledge or with other parts of the text. For assessment purposes, five errors were included in the texts to see whether students were monitoring their reading.
10.	Evaluate	Students revise and evaluate if the strategy used was useful, if the problem was resolved and if the text is now clear. Some students might evaluate if the text helped them to solve the task they had set at the beginning of the reading.

D. Appendix for Chapter 5

D.1. Reason for exclusion from reading comprehension test study

Student ID	Reasons for exclusion
6, 109, 157	Change of school
84	Expelled for behavioural problems
63, 103	Sight/learning disabilities
152	No attendance to evening workshops

E. Appendix for Chapter 6

E.1. MAI pre- and post-test responses

	Pre-test				Immediate post-test			
Item	1	2	3	4	1	2	3	4
Q1	14	30	18	10	10	34	14	14
Q2	16	23	20	13	13	17	20	22
Q3	5	22	32	13	7	17	26	22
Q4	34	21	11	6	28	24	14	6
Q5	1	9	23	39	2	9	24	37
Q6	10	30	19	13	13	25	27	7
Q7	6	14	23	29	3	20	18	31
Q8	6	8	28	30	2	11	31	28
Q9	6	18	29	19	4	19	26	23
Q10	13	24	24	11	15	24	19	14
Q11	23	30	16	3	21	29	18	4
Q12	24	19	17	12	16	30	17	9
Q13	9	16	28	19	6	19	26	21
Q14	11	20	26	15	10	18	26	18
Q15	8	36	21	7	16	25	21	10
Q16	8	20	25	19	7	19	29	17
Q17	13	33	15	11	21	23	21	7
Q18	23	28	18	3	22	30	16	4
Q19	19	27	20	6	20	26	19	7
Q20	23	24	18	7	22	25	15	10
Q21	4	21	27	20	5	17	34	16
Q22	16	34	20	2	19	27	15	11
Q23	11	14	21	26	5	13	23	31
Q24	30	24	11	7	24	23	17	8
Q25	6	18	29	19	5	20	22	25
Q26	22	32	13	5	21	28	14	9
Q27	1	14	31	26	4	18	23	27
Q28	25	20	15	12	29	19	14	10
Q29	16	20	26	10	7	16	28	21
Q30	4	23	29	16	3	18	30	21
Q31	7	25	28	12	9	22	23	18
Q32	15	22	21	14	10	31	23	8
Q33	29	19	15	9	19	26	22	5
Q34	37	24	8	3	31	24	11	6
Q35	20	27	18	7	23	18	22	9
Q36	14	30	17	11	8	24	27	13
Q37	18	28	19	7	15	25	23	9
Q38	6	24	30	12	7	24	23	18
Q39	20	29	17	6	19	24	20	9
Q40	20	20	16	16	10	26	26	10
Q41	14	30	25	3	14	26	25	7
Q42	1	24	31	16	6	21	27	18
Q43	22	30	15	5	27	22	15	8
Q44	9	15	27	21	6	20	23	23
Q45F	12	8	32	20	14	11	23	24
Q46	6	29	30	7	8	32	22	10
Q47F	10	9	38	15	10	20	23	19
Q48	17	15	30	10	11	21	24	16

Control group

	Pre-test				Immediate post-test			
Item	1	2	3	4	1	2	3	4
Q1	3	38	22	32	15	57	19	4
Q2	11	23	39	22	24	27	31	13
Q3	8	22	36	29	12	26	37	20
Q4	25	33	20	17	41	29	18	7
Q5	2	9	29	55	4	13	36	42
Q6	14	33	32	16	13	52	19	11
Q7	2	13	37	43	12	26	29	28
Q8	5	16	35	39	9	21	30	35
Q9	6	14	39	36	10	28	30	27
Q10	10	24	43	18	30	32	27	6
Q11	33	37	17	8	39	37	13	6
Q12	18	41	22	14	33	30	14	18
Q13	5	20	38	32	17	39	20	19
Q14	12	23	30	30	17	29	29	20
Q15	11	36	38	10	15	41	28	11
Q16	9	27	32	27	9	36	30	20
Q17	16	33	29	17	24	32	23	16
Q18	27	31	29	8	29	34	25	7
Q19	18	27	30	20	29	24	30	12
Q20	24	24	28	19	26	30	28	11
Q21	2	19	44	30	8	26	35	26
Q22	9	24	35	27	17	45	22	11
Q23	5	9	32	49	14	20	30	31
Q24	29	28	27	11	34	34	17	10
Q25	7	25	34	29	11	28	29	27
Q26	23	36	22	14	34	24	27	10
Q27	6	8	38	43	12	28	28	27
Q28	35	25	19	16	28	35	17	15
Q29	9	19	30	37	20	29	27	19
Q30	7	20	38	30	9	22	41	23
Q31	6	18	41	30	12	24	37	22
Q32	9	26	41	19	22	25	35	13
Q33	28	31	19	17	36	28	16	15
Q34	34	35	18	8	44	25	20	6
Q35	20	32	24	19	27	34	20	14
Q36	12	28	30	25	24	33	24	14
Q37	14	24	40	17	20	32	28	15
Q38	6	27	31	31	14	30	33	18
Q39	15	27	27	26	25	29	23	18
Q40	12	20	35	28	19	32	27	17
Q41	8	12	41	34	27	41	22	5
Q42	2	21	37	35	6	28	40	21
Q43	15	33	33	14	32	36	16	11
Q44	14	17	32	32	15	23	29	28
Q45F	13	22	35	25	11	14	35	35
Q46	6	32	36	21	19	36	22	18
Q47F	13	30	38	14	9	19	39	28
Q48	10	21	35	29	16	25	32	22

Strategy group

E.2. Results of factor analysis with 15 retained factors

Control group (Pre-test Varimax)

Item	Strategy		PC14	PC2	PC1	PC4	PC7	PC3	PC9	PC5	PC12	PC11	PC6	PC10	PC8	PC13	PC15
33	EXS	Use of diccionaries, enciclopedias or PC	0,73														
22	MET	Keep strategies in mind	0,66														
34	SUM	Take notes or write by the margin	0,64														
40	MI	Look for most important parts of text	0,58														
12	MI	Highlight, underline to remember	0,52														
7	RFB	Unknown phrase, read forw & backw		0,82													
27	RFB	Read words forw & backw to get meaning of words		0,76													
5	R2	Reread to realise why I don't understand		0,72													
16	PP	Recall text with own words		0,55													
15	SUM	Organise info			0,7												
6	R2	Reread to enhance comprehension			0,69												
18	EVA	think critically and assess info			0,65												
10	CONT	Use keywords to understand			0,48												
46	SUP	Review what I've understood and what			0,45												
3	PK	Think what I already know				0,7											
44	VIS	Visualise the events of the text				0,59											
32	SUP	I stop and think about what I read				0,52											
48	PK	Correlate text with real life					0,7										
47	MI	Highlight details of the text					-0,7										
36	PP	Paraphrase to understand					0,63										
2	PSRG	Quick scan before reading						0,73									
23	PRSG	Check text (title, pictures)						0,64									
29	CONT	Use of charts and pictures						0,52									
31	CONT	Keep reading to see whether it becomes clearer						0,47									
25	R2	Reread to make conclusions						0,4									
37	EVA	Talk to other people about the text							0,83								
24	SQ	Make questions I want to be answered							0,45								
26	RFB	Forw & backw to correlate ideas								0,77							
8	MON	Look back to find idea								0,62							
1	PSRG	Think reason for reading								0,39							
45	R2	Read texts once									0,8						
19	EVA	Check my predictions										0,7					
21	VIS	Visualise info to recall										0,62					
41	MET	I know when to use the strategies										0,45					
4	SQ	Make myself questions as teacher										0,44					
11	CONT	Find text structure (parts of text)											0,77				
39	INF	Think what will happen next or author's conclusions											0,6				
42	EVA	Think what I've learnt												0,76			
9	SCH	Adjust speed to match text & goal												-0,5			
43	PK	Correlate text with other subjects in												0,49			
17	EVA	Content matches my goal												0,4			
13	CONT	Use font and bold to identify important info													0,76		
35	SUM	Summarise to choose important information													-0,4		
20	INF	Guess what the text is about														0,72	
14	EXS	Ask teacher and friends for help														0,52	
28	SCH	I choose what to read and what to ignore														0,44	
30	CONT	Understand word meanings using clues in text															0,81
38	EVA	Review when found mismatched sentences															0,41

Control group (Post-test Varimax)

Item	Strategy		PC1	PC2	PC14	PC3	PC9	PC12	PC7	PC11	PC5	PC15	PC13	PC8	PC10	PC6	PC4
22	MET	Keep strategies in mind	0.81														
11	CONT	Find text structure (parts of text)	0.74														
24	SQ	Make questions I want to be answered	0.67														
18	EVA	think critically and assess info	0.66														
26	RFB	Forw & backw to correlate ideas	0.56														
37	EVA	Talk to other people about the text	0.54														
15	SUM	Organise info	0.49														
1	PSRG	Think reason for reading	0.42														
8	MON	Look back to find idea		0.82													
27	RFB	Read words forw & backw to get meaning of words		0.67													
38	EVA	Review when found mismatched sentences		0.5													
7	RFB	Unknown phrase, read forw & backw		0.49													
17	EVA	Content matches my goal		0.47													
5	R2	Reread to realise why I don't understand		0.44													
3	PK	Think what I already know		0.43													
30	CONT	Understand word meanings using clues in text			0.72												
19	EVA	Check my predictions			0.6												
21	VIS	Visualise info to recall			0.56												
39	INF	Think what will happen next or author's conclusions			0.41												
9	SCH	Adjust speed to match text & goal				0.64											
23	PRSG	Check text (title, pictures)				0.57											
29	CONT	Use of charts and pictures				0.53											
16	PP	Recall text with own words				0.48											
43	PK	Correlate text with other subjects in school				0.36											
6	R2	Reread to enhance comprehension					0.73										
40	MI	Look for most important parts of text					0.7										
12	MI	Highlight, underline to remember					0.48										
46	SUP	Review what I've understood and what not					0.45										
32	SUP	I stop and think about what I read					0.42										
48	PK	Correlate text with real life						0.81									
44	VIS	Visualise the events of the text						0.47									
36	PP	Paraphrase to understand							0.81								
10	CONT	Use keywords to understand							0.48								
34	SUM	Take notes or write by the margin								0.79							
41	MET	I know when to use the strategies								0.56							
47	MI	Highlight details of the text								-0.54							
35	SUM	Summarise to choose important information								0.5							
45	R2	Read texts once									0.87						
28	SCH	I choose what to read and what to ignore									-0.72						
25	R2	Reread to make conclusions									0.54						
4	SQ	Make myself questions as teacher										0.71					
20	INF	Guess what the text is about										0.58					
33	EXS	use of diccionaries, enciclopedias or PC											0.83				
31	CONT	Keep reading to see whether it becomes clearer												0.83			
13	CONT	Use font and bold to identify important info													0.78		
2	PSRG	Quick scan before reading														0.79	
14	EXS	Ask teacher and friends for help															0.71
42	EVA	Think what I've learnt															0.47

E. Appendix for Chapter 6

Control group (Post-test Varimax)

Item	Strategy		PC1	PC2	PC14	PC3	PC9	PC12	PC7	PC11	PC5	PC15	PC13	PC8	PC10	PC6	PC4
22	MET	Keep strategies in mind	0.81														
11	CONT	Find text structure (parts of text)	0.74														
24	SQ	Make questions I want to be answered	0.67														
18	EVA	Think critically and assess info	0.66														
26	RFB	Forw & backw to correlate ideas	0.56														
37	EVA	Talk to other people about the text	0.54														
15	SUM	Organise info	0.49														
1	PSRG	Think reason for reading	0.42														
8	MON	Look back to find idea		0.82													
27	RFB	Read words forw & backw to get meaning of words		0.67													
38	EVA	Review when found mismatched sentences		0.5													
7	RFB	Unknown phrase, read forw & backw		0.49													
17	EVA	Content matches my goal		0.47													
5	R2	Reread to realise why I don't understand		0.44													
3	PK	Think what I already know		0.43													
30	CONT	Understand word meanings using clues in text			0.72												
19	EVA	Check my predictions			0.6												
21	VIS	Visualise info to recall			0.56												
39	INF	Think what will happen next or author's conclusions			0.41												
9	SCH	Adjust speed to match text & goal				0.64											
23	PRSG	Check text (title, pictures)				0.57											
29	CONT	Use of charts and pictures				0.53											
16	PP	Recall text with own words				0.48											
43	PK	Correlate text with other subjects in school				0.36											
6	R2	Reread to enhance comprehension					0.73										
40	MI	Look for most important parts of text					0.7										
12	MI	Highlight, underline to remember					0.48										
46	SUP	Review what I've understood and what not					0.45										
32	SUP	I stop and think about what I read					0.42										
48	PK	Correlate text with real life						0.81									
44	VIS	Visualise the events of the text						0.47									
36	PP	Paraphrase to understand							0.81								
10	CONT	Use keywords to understand							0.48								
34	SUM	Take notes or write by the margin								0.79							
41	MET	I know when to use the strategies								0.56							
47	MI	Highlight details of the text								-0.54							
35	SUM	Summarise to choose important information								0.5							
45	R2	Read texts once									0.87						
28	SCH	I choose what to read and what to ignore									-0.72						
25	R2	Reread to make conclusions									0.54						
4	SQ	Make myself questions as teacher										0.71					
20	INF	Guess what the text is about										0.58					
33	EXS	Use of diccionaries, enciclopedias or PC											0.83				
31	CONT	Keep reading to see whether it becomes clearer												0.83			
13	CONT	Use font and bold to identify important info													0.78		
2	PSRG	Quick scan before reading														0.79	
14	EXS	Ask teacher and friends for help															0.71
42	EVA	Think what I've learnt															0.47

			Strategy group (Pre-test Varimax)														
Item	Strategy		PC1	PC3	PC2	PC9	PC11	PC5	PC12	PC13	PC10	PC6	PC7	PC14	PC8	PC4	PC15
18	EVA	Think critically and assess info	0.65														
46	SUP	Review what I've understood and what not	0.63														
36	PP	Paraphrase to understand	0.62														
16	PP	Recall text with own words	0.61														
42	EVA	Think what I've learnt	0.61														
48	PK	Correlate text with real life	0.58														
47	MI	Highlight details of the text	-0.49														
9	SCH	Adjust speed to match text & goal		0.68													
8	MON	Look back to find idea		0.61													
38	EVA	Review when found mismatched sentences		0.47													
2	PSRG	Quick scan before reading		0.44													
35	SUM	Summarise to choose important information			0.82												
12	MI	Highlight, underline to remember			0.66												
40	MI	Look for most important parts of			0.41												
41	MET	I know when to use the strategies				0.7											
33	EXS	Use of diccionaries, enciclopedias or				0.58											
34	SUM	Take notes or write by the margin				0.56											
25	R2	Reread to make conclusions				0.51											
30	CONT	Understand word meanings using clues in text					0.76										
27	RFB	Read words forw & backw to get meaning of words					0.58										
7	RFB	Unknown phrase, read forw & backw					0.57										
22	MET	Keep strategies in mind					0.53										
24	SQ	Make questions I want to be						0.69									
39	INF	Think what will happen next or author's conclusions						0.6									
19	EVA	Check my predictions							0.75								
20	INF	Guess what the text is about							0.58								
37	EVA	Talk to other people about the text							0.55								
23	PRSG	Check text (title, pictures)								0.64							
31	CONT	Keep reading to see whether it becomes clearer								0.61							
5	R2	Reread to realise why I don't								0.57							
11	CONT	Find text structure (parts of text)									0.74						
1	PSRG	Think reason for reading									0.53						
17	EVA	Content matches my goal									0.44						
45	R2	Read texts once										0.77					
28	SCH	I choose what to read and what to ignore										-0.73					
10	CONT	Use keywords to understand											0.72				
3	PK	Think what I already know											0.56				
29	CONT	Use of charts and pictures											0.43				
6	R2	Reread to enhance comprehension												0.73			
44	VIS	Visualise the events of the text													0.62		
21	VIS	Visualise info to recall													0.61		
26	RFB	Forw & backw to correlate ideas													-0.49		
13	CONT	Use font and bold to identify important info														0.76	
43	PK	Correlate text with other subjects in school														0.5	
4	SQ	Make myself questions as teacher														0.5	
32	SUP	I stop and think about what I read														-0.44	
14	EXS	Ask teacher and friends for help															0.77
15	SUM	Organise info															0.44

E. Appendix for Chapter 6

			Strategy group (Post-test Varimax)														
Item	Strategy		PC2	PC11	PC10	PC3	PC4	PC6	PC9	PC14	PC15	PC7	PC1	PC12	PC5	PC8	PC13
43	PK	Correlate text with other subjects in school	0.72														
15	SUM	Organise info	0.69														
35	SUM	Summarise to choose important information	0.65														
12	MI	Highlight, underline to remember	0.64														
30	CONT	Understand word meanings using clues in text	0.46														
36	PP	Paraphrase to understand	0.42														
23	PRSG	Check text (title, pictures)		0.67													
22	MET	Keep strategies in mind		0.67													
27	RFB	Read words forw & backw to get meaning of words		0.62													
2	PSRG	Quick scan before reading		0.6													
24	SQ	Make questions I want to be answered		0.48													
7	RFB	Unknown phrase, read forw & backw		0.44													
42	EVA	Think what I've learnt			0.76												
16	PP	Recall text with own words			0.73												
21	VIS	Visualise info to recall			0.52												
1	PSRG	Think reason for reading			0.47												
5	R2	Reread to realise why I don't understand				0.82											
6	R2	Reread to enhance comprehension				0.73											
8	MON	Look back to find idea				0.59											
3	PK	Think what I already know				0.51											
25	R2	Reread to make conclusions				0.46											
13	CONT	Use font and bold to identify important info					0.73										
9	SCH	Adjust speed to match text & goal					0.72										
20	INF	Guess what the text is about					0.66										
39	INF	Think what will happen next or author's conclusions					0.5										
19	EVA	Check my predictions					0.42										
33	EXS	Use of diccionaries, enciclopedias or PC						0.76									
11	CONT	Find text structure (parts of text)						0.6									
34	SUM	Take notes or write by the margin						0.59									
26	RFB	Forw & backw to correlate ideas							0.77								
18	EVA	think critically and assess info							0.39								
47	MI	Highlight details of the text							-0.39								
4	SQ	Make myself questions as teacher								0.82							
38	EVA	Review when found mismatched sentences									0.6						
37	EVA	Talk to other people about the text									0.53						
40	MI	Look for most important parts of text									0.46						
46	SUP	Review what I've understood and what not									0.44						
48	PK	Correlate text with real life										-0.8					
10	CONT	Use keywords to understand										0.44					
17	EVA	Content matches my goal											0.7				
31	CONT	Keep reading to see whether it becomes clearer												0.58			
41	MET	I know when to use the strategies												0.49			
44	VIS	Visualise the events of the text												0.46			
32	SUP	I stop and think about what I read												-0.44			
28	SCH	I choose what to read and what to ignore													0.81		
29	CONT	Use of charts and pictures													0.46		
45	R2	Read texts once														-0.74	
14	EXS	Ask teacher and friends for help															0.81

E.3. Case-by-variable matrix

	Por que leer		Reviso Rapido		Locate Scan		D que se trata		Repaso e imagino		Rereading to understand		Reading Forw & Back		Speed change		Use of context		Summarise		Metacognitive Strategies	
Student ID	Pret	Post	Pret	Post	Pre	Post	Pret	Post	Pret	Post	Pret	Post	Pret	Post	Pret	Post	Pret	Post	Pret	Post	Pret	Post
1																						
5																					1	
8						1																
9						2					1											
11												2									3	11
21																						5
24																					3	9
29																			1			
32																					1	
39					2																	
40																					2	
43												1										7
44					1							1									2	3
47																						1
48					2							1									3	2
51					2																5	
55												1									1	4
58																						4
61										1												8
76				2		2																3
78				1						1		1		1				1		3	3	2
79														1				1			2	8
80					1																	
81																						
83						2								2				1			2	8
87						1						1		1						2		
88				2								1						1			1	9
91									1			1						2			1	3
92		1		1	1	2				2		1		1								2
97										2												
98										1				1						1		2
101				2	1	1								1		1				1		1
105		1				1						1									2	
106					1	1				2	1	1		2								1
116																						
118		1			2									3								5
119		1				1				4		1				1						
120																					1	
124					1	1				2				3						1	2	8
125																						
126																					1	
129																					2	
133																						
134																						
136																						
137					1																	
146				1						1		1		1				1				2
147						2				3		2										
148				1		2				2		1		1		1						
153						2				3										1	2	5
154		1		1		1				1				1								2
155				2						2		2						2				7
156				2	1																	2
162						1																
164						3				1												2
165				1				1		2		4										7
170				1		1				2				1								2
171						1						1									1	5
174						2										1				1		1

References

Abbott, S. and V. Berninger (1999). 'It's never too late to remediate: Teaching word recognition to students with reading disabilities in grades 4-7'. In: *Annals of Dyslexia* 49.1, pp. 221–250.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Afflerbach, P., P. D. Pearson and S. G. Paris (2008). 'Clarifying Differences Between Reading Skills and Reading Strategies'. In: *The Reading Teacher* 61.5, pp. 364–373.

Agresti, A. (2007). *An introdution to categorical data analysis*. 2nd ed. Hoboken, NJ: Wiley-Interscience.

Akyel, A. and G. Erçetin (2009). 'Hypermedia reading strategies employed by advanced learners of English'. In: *System* 37.1, pp. 136–152.

Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

Anderson, J. R. (2009). *Cognitive Psychology and Its Implications*. 7th ed. New York: Worth Publishers.

Anderson, R. C. et al. (1985). *Becoming a Nation of Readers: The Report of the Commission on Reading*. Tech. rep. University of Illinois.

Au, K. and T. Raphael (1998). 'Curriculum and teaching in literature-based programs'. In: *Literature-based instruction: Reshaping the curriculum*. Ed. by T. Raphael and K. H. Au. Norwood, MA: Christopher Gordon, pp. 123–148.

Baddeley, A. (1992). 'Working memory'. In: *Science* 255.5044, pp. 556–559.

Baddeley, A. (1979). 'Working memory and reading'. In: *Processing of visible language* 13, pp. 355–370.

Baker, L. and B. L. Carter (2008). 'Metacognitive processes and reading comprehension'. In: *Handbook of research on reading comprehension*. Ed. by S. E. Israel and G. D. Duffy. New York: Routledge, pp. 373–388.

Baldwin, R., Z. Peleg-Bruckner and A. McClintock (1985). 'Effects of topic interest and prior knowledge on reading comprehension'. In: *Reading Research Quarterly* 20.4, pp. 497–504.

Bean, T. W. (2000). 'Reading in the Content Areas: Social Constructivist Dimensions'. In: *Handbook of Reading Research*. Ed. by M. L. Kamil et al. Vol. 3. Mahwah, NJ: Erlbaum, pp. 629–644.

Beck, I., M. G. McKeown and L. Kucan (2004). 'Direct and rich vocabulary instruction'. In: *Vocabulary instruction: Research to practice*. Ed. by J. F. Baumann and E. J. Kame'enui. Vol. 13. New York: Guilford Press, pp. 13–27.

Beck, I., C. Perfetti and M. G. McKeown (1982). 'Effects of long-term vocabulary instruction on lexical access and reading comprehension'. In: *Journal of Educational Psychology* 74, pp. 506–521.

Bernard, H. and G. Ryan (2010). *Analyzing qualitative data: Systematic approaches*. Los Angeles: SAGE Publications.

Blachowicz, C. and P. Fisher (2009). *Teaching vocabulary in all classrooms*. 4. Pearson.

Block, C. C. (2004). *Teaching comprehension: The comprehension process approach*. Allyn & Bacon.

Block, C. C. and G. D. Duffy (2008). 'Research on teaching comprehension: where we've been and where we're going'. In: *Comprehension instruction: Research-based best practices*. Ed. by C. C. Block and S. R. Parris. 2nd ed. New York: The Guilford Press, pp. 19–37.

Boulware-Gooden, R. et al. (2007). 'Instruction of Metacognitive Strategies Enhances Reading Comprehension and Vocabulary Achievement of Third-Grade Students'. In: *The Reading Teacher* 61.1, pp. 70–77.

Bowles, M. and R. Leow (2005). 'Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation'. In: *Studies in Second Language Acquisition* 27.3, pp. 415–440.

Brown, L. A. and J. Day (1983). 'Macrorules for summarizing texts: The development of expertise'. In: *Journal of Verbal Learning and Verbal Behavior* 22, pp. 1–14.

Brown, R. et al. (1996). 'A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers'. In: *Journal of Educational Psychology* 88.1, pp. 18–37.

Bruner, J. (1990). *Making sense: The child's construction of the world*. London: Routledge.

– (2006). *In search of pedagogy: The selected works of Jerome S Bruner*. London: Routledge.

Cain, K., J. Oakhill and P. Bryant (2004). 'Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills'. In: *Journal of Educational Psychology* 96.1, pp. 31–42.

Campione, J. C. (1981). 'Learning, academic achievement, and instruction'. In: *Second Annual Conference on Reading Research on the Study of Reading*.

Carrell, P. L. (1989). 'Metacognitive awareness and second language reading'. In: *The Modern Language Journal* 73, pp. 121–134.

Carrillo López, C. F., P. d. C. Gil Chaveznava and M. Morales López (2009). *Guía 11 - Preparación para el ingreso a bachillerato*. 8th ed. Santillana.

Carroll, J. M. and M. J. Snowling (2004). 'Language and phonological skills in children at high risk of reading difficulties'. In: *Journal of Child Psychology and Psychiatr* 45.3, pp. 631–640.

Carroll, J. M., C. Bowyer-Crane et al. (2011). *Developing Language and Literacy: Effective Intervention in the Early Years*. Chichester: Wiley-Blackwell.

Chamot, A. and J. O'Malley (1987). 'The cognitive academic language learning approach: A bridge to the mainstream'. In: *TESOL quarterly* 21.2, pp. 227–249.

Clark, H. and S. Haviland (1977). 'Comprehension and the given-new contract'. In: *Discourse Production and Comprehension.* 1, pp. 1–40.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Psychology Press.

Cohen, L., L. Manion and K. Morrison (2011). *Research methods in education*. 7th ed. London: Routledge.

Cook, T. D. and D. T. Campbell (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston, MA: Houghton Mifflin.

Creswell, J. W. (2003). *Research design: qualitative, quantitative, and mixed method approaches*. 2nd ed. London: Sage.

Cubukcu, F. (2008). 'Enhancing vocabulary development and reading comprehension through metacognitive strategies'. In: *Issues in Educational Research* 18.1, pp. 1–11.

Cunningham, A. (1990). 'Explicit versus implicit instruction in phonemic awareness'. In: *Journal of Experimental Child Psychology* 50.3, pp. 429–444.

Daneman, M. and P. A. Carpenter (1980). 'Individual differences in working memory and reading'. In: *Journal of Verbal Learning and Verbal Behavior* 19.4, pp. 450–466.

Davies, M. (2002). *Corpus del Español (100 million words, 1200s -1900s)*. URL: `http://www.corpusdelespanol.org`.

Davis, F. B. (1944). 'Fundamental factors of comprehension in reading'. In: *Psychometrika* 9.3, pp. 185–197.

Dillon, D. et al. (2011). 'Professional development and teacher education fo reading instruction'. In: *Handbook of Reading Research*. Ed. by M. Kamil et al. Vol. 4. New York: Routledge, pp. 629–660.

Driscoll, M. P. (2000). *Psychology of learning for instruction*. Boston, MA: Allyn & Bacon.

Duffy, G. D. and L. R. Roehler (1989). 'Why strategy instruction is so difficult and what we need to do about it'. In: *Cognitive strategy research: From basic research to educational applications*. Ed. by C. B. McCormick, G. E. Miller and M. Pressley. New York: Springer, pp. 133–154.

Duffy, G. D. (2002). 'The case for direct explanation of strategies'. In: *Comprehension instruction: Research-based best practices*. Ed. by C. C. Block and M. Pressley. New York: Guilford Press, pp. 28–41.

Durkin, D. (1978). 'What classroom observations reveal about reading comprehension instruction'. In: *Reading Research Quarterly* 14.4, pp. 481–533.

Fernández Huerta, J. (1959). 'Medidas sencillas de lecturabilidad'. In: *Consigna* 214, pp. 29–32.

Fernández Lozano, M. P., V. Jiménez Rodríguez and J. M. Alvarado Izquierdo (2010). 'La escala de conciencia lectora (ESCOLA): un instrumento para evaluar metacognición y funciones ejecutivas en tareas de lectura'. In: *Revista neuropsicología, neuropsiquiatría yneurociencias* 10.1, pp. 95–116.

Field, A. (2009). *Discovering Statistics Using SPSS*. 3rd ed. London: SAGE Publications.

Field, A. and G. Hole (2003). *How to Design and Report Experiments*. London: SAGE Publications.

Field, J. (2003). *Psycholinguistics: a resource book for students*. London: Routledge.

Flesch, R. (1948). 'A new readability yardstick'. In: *Journal of Applied Psychology* 32, pp. 221–233.

Fuchs, D. and L. Fuchs (1997). 'Peer-Assisted Learning Strategies: Making Classrooms More Responsive to Diversity'. In: *American Educational Research Journal* 34.1, pp. 174–206.

Garner, R. (1987). *Metacognition and Reading Comprehension*. Norwood, NJ: Ablex Publication Corp.

Garner, R. and K. Kraus (1981). 'Good and poor comprehender differences in knowing and regulating reading behaviors'. In: *Educational Research Quarterly* 6, pp. 5–12.

Gass, C. S., Y. I. Demsky and P. Cameron Martin (1998). 'Factor analysis of the WISC-R (Spanish version) at 11 age levels between 6 1/2 and 16 1/2 years'. In: *Journal of Clinical Psychology* 54.1, pp. 109–113.

Gersten, R. and D. Carnine (1986). 'Direct Instruction in Reading Comprehension'. In: *Educational Leadership* 43.7, pp. 70–78.

Gillies, R. and A. Khan (2009). 'Promoting reasoned argumentation, problem-solving and learning during small-group work'. In: *Cambridge Journal of Education* 39.1, pp. 7–27.

Grabe, W. (2008). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Pressd.

Grabe, W. and F. Stoller (2002). *Teaching and researching reading*. Harlow: Longman.

Guía para Bachillerato (n.d.).

Hacker, D. (1998). 'Definitions and empirical foundations'. In: *Metacognition in Educational Theory and Practice*. Ed. by D. J. Hacker, J. Dunlosky and A. Graesser. Mahwah, NJ: Lawrence Erlbaum, pp. 1–24.

Hakkarainen, A., L. Holopainen and H. Savolainen (2013). 'Mathematical and reading difficulties as predictors of school achievement and transition to secondary education'. In: *Scandinavian Journal of Educational Research* 57.5, pp. 488–506.

Harley, T. (2001). *The psychology of language. From data to theory*. Ed. by 2. Hove: Psychology Press.

Harris, K., P. Alexander and S. Graham (2008). 'Michael Pressley's Contributions to the History and Future of Strategies Research'. In: *Educational Psychologist* 43.2, pp. 86–96.

Hoover, W. and P. Gough (1990). 'The simple view of reading'. In: *Reading and writing* 2.2, pp. 127–160.

Houtveen, A. A. M. and W. J. C. M. van de Grift (2007). 'Effects of metacognitive strategy instruction and instruction time on reading comprehension'. In: *School Effectiveness and School Improvement* 18.2, pp. 173–190.

Howe, K. R. (1988). 'Against the quantitative-qualitative incompatibility thesis, or dogmas die hard'. In: *Educational Researcher* 17.8, pp. 10–16.

Hulme, C. and M. J. Snowling (2005). 'Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart'. In: *Scientific Studies of Reading* 9.4, pp. 351–365.

Inchausti de Jou, G. and T. M. Sperb (2009). 'Lectura comprensiva: Un estudio de intervención'. In: *Interamerican Journal of Psychology* 43.1, pp. 12–21.

INEE (2006a). *La calidad de la educación básica ayer, hoy y mañana*. Instituto Nacional para la Evaluación de la Educación.

– (2006b). *La calidad de la educación básica en México*. Instituto Nacional para la Evaluación de la Educación.

Instituto Nacional de Estadística y Geografía (2011). *Sistema nacional de clasificación de ocupaciones 2011*. Instituto Nacional de Estadística y Geografía (México).

Jacobs, J. and S. G. Paris (1987). 'Children's metacognition about reading: Issues in definition, measurement, and instruction'. In: *Educational Psychologist* 22.3-4, pp. 255–278.

Jiménez Rodríguez, V. (2004). 'Metacognición y Comprensión de la Lectura: Evaluacion de los Componentes Estratégicos (Procesos y Variables) Mediante la Elaboración de un Escala de Conciencia Lectora (ESCOLA)'. PhD thesis. Universidad Complutense de Madrid.

Jiménez Rodríguez, V. et al. (2009). 'La medición de las estrategias metacognitivas mediante la escala de conciencia lectora ESCOLA'. In: *Revista de Investigación Psicoeducativa* 7.2, pp. 185–205.

Johnston, P. (1984). 'Prior knowledge and reading comprehension test bias'. In: *Reading Research Quarterly* 19.2, pp. 219–239.

Juel, C., P. Griffith and P. Gough (1986). 'Acquisition of literacy: A longitudinal study of children in first and second grade'. In: *Journal of Educational Psychology* 78.4, pp. 243–255.

Just, M. A. and P. A. Carpenter (1980). 'A theory of reading: From eye fixations to comprehension'. In: *Psychological Review* 87, pp. 329–354.

Kaiser, H. (1960). 'The application of electronic computers to factor analysis'. In: *Education and Psycological Measurement* 20, pp. 141–151.

Kintsch, W. (1988). 'The role of knowledge in discourse comprehension: A constructive integration model'. In: *Psychological Review* 95, pp. 163–182.

Kintsch, W. and A. K. Rawson (2005). 'Comprehension'. In: *The Science of Reading: A Handbook*. Ed. by M. J. Snowling and C. Hulme. Oxford: Blackwell Publishing, pp. 209–226.

Kintsch, W. and T. A. Van Dijk (1978). 'Toward a model of text comprehension and production'. In: *Psychological Review* 85, pp. 363–394.

LaBerge, D. and S. Samuels (1974). 'Toward a theory of automatic information processing in reading'. In: *Cognitive Psychology* 6.2, pp. 283–323.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

Lehto, J. et al. (2001). 'National survey of reading comprehension in Finland'. In: *Journal of Research in Reading* 24.1, pp. 99–110.

Levin, J. R. and M. Pressley (1981). 'Improving children's prose comprehension: Selected strategies that seem to succeed'. In: *Children's prose comprehension: Research and practice*. Ed. by C. M. Santa and B. L. Hayes. Newark: International Reading Association, pp. 44–71.

Lundberg, I., J. Frost and O. Petersen (1988). 'Effects of an extensive program for stimulating phonological awareness in preschool children.' In: *Reading Research Quarterly* 23.3, pp. 263–284.

Madariaga Orbea, J. and E. Martínez Villabeitia (2010). 'The teaching of reading comprehension and metacomprehension strategies . A program implemented by teaching staff Introduction: The difficulties of text comprehension in the classroom'. In: *Anales de Psicología* 26, pp. 112–122.

Markman, E. M. (1979). 'Realizing that you don't understand: Elementary school children's awareness of inconsistencies'. In: *Child development*, pp. 643–655.

Marsden, E. and C. Torgerson (2012). 'Single group, pre- and post-test research designs: Some methodological concerns'. In: *Oxford Review of Education* 38.5, pp. 583–616.

Martinez Rizo, F. (2008). *¿Avanza o retrocede la calidad educativa? Tendencias y perspectivas de la educación básica en México*.

Martínez, N. (Dec. 2013). 'Tras resultados de PRUEB PISA, habrá cambios: SEP'. In: *El Universal*. URL: http://www.eluniversal.com.mx/nacion-mexico/2013/tras-resultados-de-prueba-pisa-habra-cambios-sep-970200.html.

Mathes, P. et al. (1998). 'Peer-assisted learning strategies for first-grade readers: Responding to the needs of diverse learners'. In: *Reading Research Quarterly* 33.1, pp. 62–94.

Mathews, M. (1966). *Teaching to read: Historically considered*. New York: John Wiley.

Mayer, R. (1988). 'Learning strategies: An overview'. In: *Learning and study strategies*. Ed. by C. Weinstein, E. T. Goetz and P. A. Alexander. New York: Academic Press.

McGowan, H. M. (2011). 'Planning a Comparative Experiment in Educational Settings'. In: *Journal of Statistics Education* 19.2, pp. 1–19.

McLain, K., B. Gridley and D. McIntosh (1991). 'Value of a scale used to measure metacognitive reading awareness'. In: *The Journal of Educational Researc* 85.2, pp. 81–87.

McNamara, T. P., D. L. Miller and J. D. Bransford (1991). 'Mental models and reading comprehension'. In: *Handbook of Reading Research*. Ed. by M. L. Kamil, P. Mosenthal and P. D. Pearson. Vol. 2. New York: White Plains, Longman, pp. 490–511.

Miholic, V. (1994). 'An Inventory to Pique Students' Metacognitive Awareness of Reading Strategies'. In: *Journal of Reading* 38.2, pp. 84–86.

Mitchell, R. and F. Myles (2004). *Second language learning theories*. 2nd ed. London: Hodder Education.

Mitchell, R., F. Myles and E. Marsden (2013). *Second language learning theories*. 3rd ed. Abingdon: Routledge.

Mokhtari, K. and C. A. Reichard (2002). 'Assessing Students' Metacognitive Awareness of Reading Strategies'. In: *Journal of Educational Psychology* 94.2, pp. 249–259.

Mokhtari, K. and R. Sheorey (2008). *Reading strategies of first- and second-language learners: see how they read*. Norwood, MA: Christopher-Gordon Publishers.

Moore, L., A. Graham and I. Diamond (2003). 'On the feasibility of conducting randomised trials in education: case study of a sex education intervention'. In: *British Educational Research Journal* 29.5, pp. 673–689.

Morse, J. (2003). 'Principles of Mixed and Multi Method Research Design'. In: *Handbook of Mixed Methods in Social and Behavioral Research*. Ed. by A. Tashakkori and C. Teddlie. London: SAGE Publications.

Nagy, W. (1988). *Teaching Vocabulary To Improve Reading Comprehension*. Urbana, IL: National Council of Teachers of English.

Nathaniel, G. (2007). 'The paradigm wars and their aftermath'. In: *Educational research and evidence-based practice*. Ed. by M. Hammersley. London: SAGE Publications.

Nation, P. and J. Newton (1997). *Teaching vocabulary*. URL: `http://sites.google.com/site/morrowteacher/elementary-me/NationTeachingVocab.pdf`.

Neufeld, B. and D. Roper (2003). *Coaching: A strategy for developing instructional capacity: promises and practicalities*. Aspen Institute.

Nisbet, J. and J. Shucksmith (1986). *Learning strategies*. Routledge.

Norris, J. M. and L. Ortega (2006). *Synthesizing research on language learning and teaching*. Amsterdam: John Benjamins.

OECD (2001). *Knowledge and skills for life*. Paris: OECD Publishing.

– (2002). *Programme for International Student Assessment: Sample tasks from the 2000 assessment of reading, mathematical and scientific literacy*. Paris: OECD Publishing.

– (2006). *Education at a glance: OECD Indicators*. Paris: OECD Publishing.

– (2009). *PISA 2009 assessment framework - key competences in reading, mathematics and science*. OECD Publishing.

– (2013). *OECD PISA results database*. URL: `http://www.oecd.org/pisa/pisaproducts/`.

Paris, S. G., D. Cross and M. Y. Lipson (1984). 'Informed strategies for learning'. In: *Journal of Educational Psychology* 76, pp. 1239–1252.

Paris, S. G. and E. E. Hamilton (2008). 'The development of childrens' reading comprehension'. In: *Handbook of research on reading comprehension*. Ed. by S. E. Israel and G. D. Duffy. New York: Routledge, pp. 32–53.

Paris, S. G., M. Y. Lipson and K. K. Wixson (1983). 'Becoming a strategic reader'. In: *Contemporary Educational Psychology* 8, pp. 293–316.

Paris, S. G. and M. Myers (1981). 'Comprehension monitoring, memory, and study strategies of good and poor readers'. In: *Journal of Literacy Research* 13.1, pp. 5–22.

Pearson, P. D. (1985). 'Changing the face of reading comprehension instruction'. In: *Reading Teacher* 38, pp. 724–738.

– (2008). 'The roots of reading comprehension instruction'. In: *Handbook of research on reading comprehension*. Ed. by S. E. Israel and G. D. Duffy. New York: Routledge, pp. 3–31.

Pearson, P. D. and L. Fielding (1991). 'Comprehension instruction'. In: *Handbook of Reading Research*. Ed. by R. Barr et al. Vol. 2. New York: Longman, pp. 819–860.

Pearson, P. D. and M. Gallagher (1983). 'The instruction of reading comprehension'. In: *Contemporary Educational Psychology* 8.3, pp. 317–344.

Pereira-Laird, J. and F. Deane (1997). 'Development and validation of a self-report measure of reading strategy use'. In: *Reading Psychology* 18.3, pp. 185–235.

Perfetti, C. (2007). 'Reading ability: Lexical quality to comprehension'. In: *Scientific Studies of Reading* 11.4, pp. 357–383.

Perfetti, C., N. Landi and J. Oakhill (2005). 'The acquisition of reading comprehension skill'. In: *The Science of Reading: A Handbook*. Ed. by M. J. Snowling and C. Hulme. Oxford: Blackwell Publishing, pp. 227–247.

Phakiti, A. (2003). 'A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance'. In: *Language Testing* 20.1, pp. 26–56.

Pressley, M. (2000). 'What should comprehension instruction be the instruction of?' In: *Handbook of Reading Research*. Ed. by M. Kamil et al. Vol. 3. Mahwah, NJ: Erlbaum, pp. 545–561.

Pressley, M. and P. Afflerbach (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Erlbaum.

Pressley, M., P. B. El-Dinary et al. (1992). 'Beyond Direct Explanation: Transactional Instruction of Reading Comprehension Strategies'. In: *The Elementary School Journal* 92.5, pp. 513–513.

Pressley, M., F. Goodchild et al. (1989). 'The Challenges of Classroom Strategy Instruction'. In: *The Elementary School Journal* 89.3, pp. 301–342.

Pring, R. (2004). *Philosophy of Educational Research*. 2nd ed. London: Continuum International Publishing.

Pritchard, A. (2009). *Ways of learning: learning theories and learning styles in the classroom*. 2nd ed. London: Routledge.

Pritchard, R. (1990). 'The effects of cultural schema on reading processing strategies'. In: *Reading Research Quarterly* 25, pp. 273–295.

RAND Reading Study Group (2001). *Reading for understanding: Toward an R&D program in reading comprehension*. Tech. rep. RAND.

Raphael, T. et al. (2008). 'Approaches to teaching reading comprehension'. In: *Handbook of research on reading comprehension*. Ed. by E. S. Israel and G. D. Duffy. New York: Routledge, pp. 449–465.

Rattray, J. and M. C. Jones (2007). 'Essential elements of questionnaire design and development'. In: *Journal of Clinical Nursing* 16, pp. 234–243.

Rayner, K. and A. Pollatsek (1989). *The psychology of reading*. Prentice Hall International.

Real Academia Española (2001). *Diccionario de la lengua española*.

Robles Robles, D. and F. Olvera (2002). *Nueva guía mágica para ingreso al bachillerato*. Fernández Editores.

Rodriguez, V. L. and J. Prewitt Diaz (1990). 'Correlations among GPA and scores on the Spanish version of WISC-R and the Woodcock-Johnson Achievement subtests for 10- to 12-year-old Puerto Rican children'. In: *Psychological Reports* 66.2, pp. 563–566.

Rosenshine, B. and C. Meister (1994). 'Reciprocal teaching: A review of the research'. In: *Review of Educational Research* 64, pp. 479–530.

Rumelhart, D. E. (1994). 'Toward an interactive model of reading'. In: *Theoretical models and processes of reading*. Ed. by R. Ruddell, M. Ruddell and H. Singer. 4th ed. Newark: International Reading Association.

Sale, J. E., L. H. Lohfeld and B. Kevin (2002). 'Revisiting the quantiative-qualitative debate: implications for mixed-methods research'. In: *Quality and quantity* 36.1, pp. 43–53.

Samuels, S. and M. L. Kamil (1984). 'Models of reading process'. In: *Handbook of Reading Research*. Ed. by P. D. Pearson et al. White Plains, NY: Longman, pp. 185–224.

Samuels, S. (1988). 'Decoding and automaticity: Helping poor readers become automatic at word recognition'. In: *Reading Teacher* 41.8, pp. 756–760.

Schmitt, M. (1990). 'A questionnaire to measure children's awareness of strategic reading processes'. In: *The Reading Teacher* 43.7, pp. 454–461.

Schunk, D. (2000). *Learning theories: An educational perspective*. 3rd ed. Upper Saddle River, N.J: Merrill Prentice Hall.

SEP (2012). *Programa Nacional de Lectura*. URL: `http://www.lectura.dgme.sep.gob.mx/11mas5_2012/`.

Shadish, W. R., T. D. Cook and D. T. Campbell (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Simpson, M. L. and S. L. Nist (2000). 'An update on strategic learning: It's more than textbook reading strategies'. In: *Journal of Adolescent & Adult Literacy* 43.6, pp. 528–541.

Singhal, M. (2001). 'Reading proficiency, reading strategies, metacognitive awareness and L2 readers'. In: *The Reading Matrix* 1, pp. 1–23.

Smidt, S. (2011). *Introducing Bruner: A guide for practitioners and students in early years education*. London: Routledge.

Smith, L. (1997). *Piaget, Vygotsky and beyond: Future issues for developmental psychology and education*. London: Routledge.

Spiro, R. J. et al. (1987). 'Cognitive flexibility and transfer in complex content domains'. In: *Executive control processes in reading*, pp. 177–199.

Stanovich, K. (1980). 'Toward an interactive-compensatory model of individual differences in the develpment of reading fluency'. In: *Reading Research Quarterly* 16, pp. 32–71.

Tan, A. and T. Nicholson (1997). 'Flashcards revisited: Training poor readers to read words faster improves their comprehension of text'. In: *Journal of Educational Psychology* 89.2, pp. 276–288.

Tashakkori, A. and C. Teddlie (1998). *Mixed methodology: combining qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE.

– (2003). *Handbook of Mixed Methods in Social and Behavioral Research*. London: SAGE Publications.

Thomas, K. and M. A. Barksdale-Ladd (2000). 'Metacognitive Processes: Teaching Strategies in Literacy Education Courses'. In: *Reading Psychology* 21.1, pp. 67–84.

Tierney, R. J. and J. W. Cunningham (1980). *Research on teaching reading comprehension*. Tech. rep. University of Illinois.

– (1984). 'Research on teaching reading comprehension'. In: *Handbook of Reading Research*. Ed. by P. D. Pearson et al. White Plains, NY: Longman, pp. 609–655.

Tierney, R. J., J. E. Readence and E. K. Dishner (2005). *Reading strategies and practices: A compendium*. 6th ed. Boston: Allyn & Bacon.

Torgerson, C. and D. Torgerson (2003). 'The design and conduct of randomised controlled trials in education: lessons from health care'. In: *Oxford Review of Education* 29.1, pp. 67–80.

Torgerson, D. and C. Torgerson (2008). *Designing randomised trials in health, education and the social sciences : an introduction*. Basingstoke: Palgrave Macmillan.

Treviño, E. P. et al. (2006). *Prácticas docentes para el desarrolo de la comprensión lectora de primeria*. Instituto Nacional para la Evaluación de la Educación.

UNESCO (2004). *Education for all - The quality imperative*.

– (2012). *International Standard Classification of Education - ISCED 2011*. Tech. rep. UNESCO Institute for Statistics.

Van den Broek, P. et al. (2011). 'When a reader meets a text: the role of standards of coherence in reading comprehension'. In: *Text relevance and learning from text*. Ed. by M. McCrudden, J. Magliano and G. Schraw. Greenwich, CT: Information Age Publishing, pp. 123–140.

Van Dijk, T. A. and W. Kintsch (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children - Revised*. New York: Psychological Corporation.

Weinberg, S. L. and S. K. Abramowitz (2002). *Data Analysis for the Behavioral Sciences Using SPSS*. Cambridge: Cambridge University Press.

Werlich, E. (1976). *A text grammar of English*. Heidelberg: Quelle und Meyer.

Wilawan, S. (2007). *Lexical cohesion and metacognitive strategy training: An integrated approach to main idea comprehension*. Leeds Working Papers in Linguistics and Phonetics.

Young, K. A. (2005). 'Direct from the source: the value of 'think-aloud' data in understanding learning'. In: *Journal of Educational Enquiry* 6.1, pp. 19–33.

Yuill, N. and J. Oakhill (1988). 'Effects of inference awareness training on poor reading comprehension'. In: *Applied Cognitive Psychology* 2.1, pp. 33–45.

– (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge: Cambridge University Press.

Zinar, S. (2000). 'The Relative Contributions of Word Identification Skill and Comprehension-Monitoring Behavior to Reading Comprehension Ability'. In: *Contemporary Educational Psychology* 25.4, pp. 363–377.