Reading Comprehension and Working Memory's Executive Processes: An Intervention Study in Primary School Students

Juan A. García-Madruga
María Rosa Elosúa
Universidad Nacional de Educación a Distancia, Madrid, Spain

Laura Gil
University of Valencia, Spain

Isabel Gómez-Veiga
José Óscar Vila
Isabel Orjales
Antonio Contreras
Raquel Rodríguez
Universidad Nacional de Educación a Distancia, Madrid, Spain

María Ángeles Melero
University of Cantabria, Santander, Spain

Gonzalo Duque
Universidad Nacional de Educación a Distancia, Madrid, Spain

Reading Research Quarterly, 48(2)
pp. 155–174 | doi:10.1002/rrq.44

ABSTRACT

Reading comprehension is a highly demanding task that involves the simultaneous process of extracting and constructing meaning in which working memory's executive processes play a crucial role. In this article, a training program on working memory's executive processes to improve reading comprehension is presented and empirically tested in two experiments with third-grade primary school students. Experiment 1 showed a greater gain after training the experimental group in contrast to the control group in reading comprehension and intelligence. In experiment 2, we focused on the training processes and compared training results of high and low pretest reading comprehension groups. Results confirmed the increase in reading comprehension, intelligence, and executive processes and showed that the low group reached a greater gain in reading comprehension after training than the high group did. The results of these experiments and their limitations are discussed in the context of recent theories on the role of executive processes in reading comprehension and the possibility of training working memory and intelligence.

Working memory (WM) is a central component of the cognitive neuroscience view of the human mind. WM capacity refers to the number of items that can be recalled during a complex WM task. From a conceptual perspective, there is no general agreement about the definition of WM capacity, namely, because there are diverse theories that show some basic agreement but emphasize different aspects of WM (see Miyake & Shah, 1999). Nevertheless, there is no question that one of the most influential of these theories is the multiple-component model proposed by Baddeley and Hitch (1974; Baddeley, 1986, 2000, 2007). According to this theoretical model, the WM system includes two domain-specific storage structures or slave systems (the phonological loop and the visuospatial sketchpad), an episodic buffer that links the two prior components with long-term memory, and a central executive. The central executive is the main component of the WM system. It not only has to coordinate the other components but is also in charge of the attentional control of information.

There are also two related and influential models of WM: Cowan's (1999) embedded-processes model and Engle's (2001; Unsworth & Engle, 2007) general capacity model. Unlike Baddeley's multiple-component model, these two models neglect the existence of domain-specific components in WM. On the one hand, in Cowan's model, there is just one basic memory repository (similar to long-term memory) in which information can be activated at different levels. According to Cowan, WM entails an embedded subset of activated information, which is more salient by bringing it into the focus of attention. On the other hand, Engle and colleagues define WM more explicitly as a system consisting of highly activated long-term memory traces that are active above threshold as short-term memory representational components.

In spite of their differences, Baddeley's, Cowan's, and Engle's models all share the idea of a domain-general central executive whose main functions are to focus and switch attention, to activate and update representations, and to inhibit automatic processes and discard irrelevant information (see Baddeley, 2007; Cowan, 2005; Engle, 2002; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). The executive functions of WM during reading comprehension are the main focus of this article.

WM is closely related to general intelligence as a number of studies have shown (see, e.g., Ackerman, Beier, & Boyle, 2005; Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Kyllonen & Christal, 1990). However, the exact nature of the processes that underlie and explain this relation is a matter of debate among researchers. One main piece of evidence established is the relationship between WM and fluid intelligence—that is, the individual's ability to reason with novel problems (see Cornoldi, 2006; Kane et al., 2004; Oberauer, Schulze, Wilhelm, & Süß, 2005).

According to Engle and colleagues (Engle, Tuholski, Laughlin, & Conway, 1999; Kane & Engle, 2002; Unsworth & Engle, 2005), executive control is the crucial component in the explanation of the relations between WM and fluid intelligence: the individual's ability to maintain attentional control in complex tasks that require one to resist and control interfering information. But not all of the executive functions seem to be equally related to intelligence. Friedman et al. (2006; see also Chen & Li, 2007) found that updating predicted fluid intelligence in young adults better than inhibition or switching did.

In more recent work, Belacchi, Carretti, and Cornoldi (2010) investigated the role of updating and other various WM measures in predicting fluid intelligence measured by means of the colored Raven's Progressive Matrices test in students ages 5–11 years. The results showed a strong relation between fluid intelligence and diverse measures of WM capacity and executive processes, but the best predictor of fluid intelligence was updating. These findings throw light on another focus of our article, the relationship between WM's executive processes and fluid intelligence in students.

Reading comprehension demands that people store text information recently decoded and that they apply complex processes of meaning construction to arrive at an integrated representation or situational model (e.g., Kintsch, 1998). In other words, we consider text comprehension a highly demanding cognitive task that implies the simultaneous process of extracting and constructing meaning (Snow & Sweet, 2003). To extract and construct meaning, readers must engage in a process of knowledge activation and use, which we call making inferences (Kintsch, 1998). As numerous authors have maintained, WM plays a crucial role in storing the intermediate and final products of readers' computations, as well as coordinating the processes of constructing and integrating the semantic representation from a text (e.g., Cain, 2006; Ericsson & Kintsch, 1995; Gathercole & Baddeley, 1993; Just & Carpenter, 1992).

Some authors have stressed the importance of active processing during reading to achieve this semantic representation (see Britton & Graesser, 1996; García Madruga, Martín Cordero, Luque, & Santamaría, 1992; Kintsch, 1998; van Oostendorp & Goldman, 1999). The key idea is that comprehension depends essentially on the reader's active use of knowledge that guides his or her strategies toward the construction of meaning from textual information (García Madruga, 2006). This active process of building meaning, as well as the necessary metacognitive monitoring during reading (Baker, 1989; Wagoner, 1983), underscores the importance of attentional control and enhances even more the role of executive control processes in reading comprehension.

Hence, it is unsurprising that learning to read comprehensively is often a rather complicated acquisition. It demands that the perception and identification of letters and words is automated so cognitive resources are left free to be assigned to the construction of meaning and the representation of the situation that the text describes. However, even if the superficial tasks implied in reading are adequately automated, some difficulties may appear at higher levels of comprehension (see Oakhill, & Cain, 2007).

The relationship between WM span and reading comprehension has been well established in the literature (see, e.g., Daneman & Merikle, 1996). WM capacity is closely related to diverse reading comprehension skills (e.g., De Beni, Palladino, Pazzaglia, & Cornoldi, 1998; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000). Moreover, students with high WM scores typically show good comprehension reading skills, and conversely, students with poor WM scores tend to perform below average on reading comprehension measures (see Baddeley, 2007; Cain, Oakhill, & Bryant, 2004; Swanson & Howell, 2001; Vukovic & Siegel, 2006).

For example, Cain et al. (2004) reported data from a longitudinal study that addresses the relationship between WM capacity and reading comprehension skills in students ages 8, 9, and 11 years. At each point in time, WM and component skills of comprehension predicted unique variance in reading comprehension after controlling for word reading, vocabulary, and verbal abilities. More recently, Vukovic and Siegel (2006) extended these findings by demonstrating that WM plays an important role in reading comprehension even after controlling for phonological awareness and rapid naming.

As for the involvement of the diverse WM components in reading comprehension, verbal WM is an obvious component, as shown by studies that have used Daneman and Carpenter's (1980) reading span test (RST; e.g., García-Madruga, Elosúa, Gutiérrez, Gárate, & Luque, 1999; Hannon & Daneman, 2004). RST is a good measure of verbal WM, but because it requires some kind of attentional control, RST is also a measure of WM's central executive (see Engle & Oransky, 1999; García-Madruga, Gutiérrez, Carriedo, Luzón, & Vila, 2007; Whitney, Arnett, Driver, & Budd, 2001).

Along this line, an increasing number of authors have underscored the role of the diverse yet interrelated executive processes of WM in reading comprehension. In particular, Swanson, Howard, and Saez (2006) pointed out that the executive function of coordinating cognitive operations is required by the integration of information from text and long-term memory; Palladino, Cornoldi, De Beni, and Pazzaglia (2001; see also Carretti, Cornoldi, De Beni, & Romanò, 2005) have linked WM's updating to reading comprehension skills; and De Beni and Palladino (2000; see also Carretti, Borella, Cornoldi, & De Beni, 2009) and Savage, Cornish, Manly, and Hollis (2006) have underscored the function of inhibiting and discarding information in reading comprehension.

In this article, we consider that the comprehension of difficult texts such as those read by students at school requires readers to apply all of the executive processes of WM: focusing on complex reading tasks and switching attention between diverse textual information and the required cognitive tasks, activating knowledge from long-term memory and updating an integrated representation of the meaning of the text, and inhibiting possible representations and discarding irrelevant information.

In the last decade, some authors have highlighted the importance of having the teaching of WM and its executive processes embedded into the classroom curriculum. For instance, Gathercole, Lamont, and Alloway (2006) defended the importance of identifying WM problems as a source of learning difficulty in individual students and of reducing the opportunities for learning failures by minimizing WM demands in classroom activities. As a way to achieve this goal with first and second graders, school staff received guidance to identify WM failures in the classroom, as well as instruction about how to minimize this kind of failure in individual students (e.g., cutting down the processing load of the task, using external memory aids, ensuring that the student can remember the task).

Along the same line, Meltzer, Pollica, and Barzillai (2007) developed a classroom intervention based on teaching strategies that address executive functions in the classroom. Likewise, Gaskins, Satlow, and Pressley (2007) addressed a systematic and goal-oriented approach to teach reading comprehension strategies in elementary school. Their approach promotes the use of different executive processes to enable readers to monitor whether what they read makes sense and to take charge of whether they understand what they read.

Some other authors have even shown that students of different ages, with and without attention-deficit hyperactivity disorder (ADHD), and young adults can reach a sustained enhancement on diverse WM and intelligence measures after going through an intensive adaptive WM training program (see Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi et al., 2010; Klingberg, Forssberg, & Westerberg, 2002; Thorell, Lindqvist, Nutley, Bohlin, & Klingberg, 2009). In spite of the close relationship between WM and reading comprehension, most of these latter studies on WM training did not analyze their effect on reading comprehension. In our knowledge, there is only one study (Dahlin, 2011) that showed an increase in reading comprehension after WM training. In Dahlin's work, primary school students with special needs were trained using a procedure based on Klingberg et al.'s (2002) for ADHD students. After daily individual training at school for 30–40 minutes over a period of five weeks, the results showed a substantial improvement in reading comprehension tasks.

There are, however, some recent publications that have questioned the efficacy of training on WM. Two of them are particularly relevant: Melby-Lervåg and Hulme (2012) and Shipstead, Redick and Engle (2012). The first is a meta-analytic review of thirty-three studies with clinical and typically developing samples of students and adults. Melby-Lervåg and Hulme (2012) conclude that these training programs yield only near-transfer effects and that there is no evidence that these effects are durable. Likewise, these authors cast doubt on the relevance and theoretical basis of studies seeking to train WM to enhance cognitive functioning. They claim that they "do not appear to be based on any clear theory of the processes involved or any clear task analysis" (p. 13). As Melby-Lervåg and Hulme explicitly acknowledge, the problem with meta-analyses is that they bring together studies that widely differ in their characteristics and theoretical perspectives.

As we have just claimed, some of the studies included in Melby-Lervåg and Hulme's (2012) meta-analysis are relevant toward the enhancement of cognitive functioning and also are theoretically sound. However, in agreement with Melby-Lervåg and Hulme, in some of the training programs, a clearer analysis of the processes involved and the tasks used is missing. In this article, we attempt to be more precise in the analysis of the processes involved and the tasks used in training.

The second critical paper is that of Shipstead, Redick, and Engle (2012). It undertakes a more theoretically based analysis and review of studies that focus on the training of WM. These authors pose three main general concerns of studies on WM training:

1. The use of single tasks to decide a change in one ability, such as only using the RST for WM or Raven's Progressive Matrices for fluid intelligence
2. The lack of a consistent use of valid WM tasks, different from those used in training, to evaluate WM training effects; they rightly also criticize the use of simple short-term memory tasks instead of complex WM span tasks
3. The use of noncontact control groups, in which individuals "participate in pre- and posttest sessions but are not otherwise engaged in the experiment" (p. 635). The especial involvement and higher motivation of experimental participants can affect the results of training groups.

Shipstead et al. (2012) conclude that the results found in these studies are preliminary; although they are in some respects clearly promising, they do not provide sufficient evidence of the efficacy of WM training. We also attempt in this article to avoid at least some of the above criticisms by using more appropriate WM span tasks (not short-term memory tasks), different tasks in the pre- and posttest than used in training, and more than one measure for one ability (WM's executive processes) and active contact groups, when possible.

The main aim of the present article is to train normally developing students on WM's executive processes involved in reading to improve their reading comprehension abilities. Our perspective shares many features with that of the WM training programs just discussed, but it is also partially different: We did not train participants using WM tasks (except for the case of anaphora and analogies WM tasks). Instead, we trained them with text-processing tasks that demand high attentional control—that is, reading comprehension tasks in which WM's executive processes are particularly involved (focusing, switching, connecting with prior knowledge, semantic updating in WM, and inhibition).

Experiment 1

In this experiment, we evaluated reading comprehension, WM, and intelligence in two groups of primary school students and carried out an intervention to improve reading comprehension. Thus, the aim of the study was to evaluate the impact of a training program designed to improve reading comprehension by boosting the main functions of the central executive related to it: focusing, switching, the activation of long-term representations and updating, and the inhibition of irrelevant information. We were also interested in assessing possible transfer effects of training to intelligence and WM's executive processes measures.

Our hypotheses were twofold:

1. In the experimental group, we predict a significant increase after training in the posttest measure of reading comprehension. The increase in the experimental group in reading comprehension will be significantly higher than that obtained in the control group.
2. There will be positive correlations in the pretest among the three cognitive variables studied: reading comprehension, WM, and intelligence.

Method

Design and Participants

We used an intervention design with pretest and posttest measures and an active control group. Thirty-five third-grade students from a middle socioeconomic-level school in Alcobendas (Comunidad de Madrid) participated in the experiment. Data are reported from the 31 students ages 8–9 years who completed the training program ($M = 8.42$, $SD = 0.46$). They were randomly assigned either to the experimental group ($n = 15$, $M = 8.52$, $SD = 0.49$) or the control group ($n = 16$, $M = 8.32$, $SD = 0.42$).

Procedure

The pretest evaluation was carried out at the end of grade 3 (May 2009), whereas the intervention and posttest evaluation were performed by the students during fourth grade. Students in the experimental group were trained for 12 days in their normal classroom for 50 minutes daily over a period of four weeks. They carried out a number of reading comprehension tasks directly related to one or more of the functions of the central executive. During the training process, all participants received a workbook that included the diverse exercises to be performed in each session. Participants were asked to write their responses to the diverse exercises in these workbooks, which were collected by experimenters at the end of each session. Students in the control group received normal class instruction from their teacher in Spanish language and reading comprehension instead of experimental training. All participants were assessed on measures of reading comprehension, WM, and intelligence before and after training.

Pretest and Posttest Measures

Reading Comprehension

To measure reading comprehension, we used a Spanish version (EDICOLE) of the Diagnostic Assessment of Reading Comprehension (DARC; August, Francis, Hsu, & Snow, 2006; Francis et al., 2006). This relatively new test is based on a theoretical analysis of reading comprehension and consists of four main components (Hannon & Daneman, 2001). The task requires students to silently read three short texts and answer 44 related comprehension questions. Presented in a narrative style, the texts consist of four small paragraphs that describe transitive relations among a set of real and artificial entities—for example, "Maria likes to eat fruit. Most of all, she likes to eat nuras. A nura is like an orange. But a nura is bigger than an orange."

Combining the information in the text with world knowledge should, in principle, allow for the construction of a five-entity–long linear ordering along a dimension that is likely to be familiar to all students. Three of the entities are unknown to all readers (artificial terms) and are presented as nonsense words, whereas two of the entities referred to are likely to be known by all students (real terms) and differ strikingly on the critical dimension. After each text, readers are asked a series of 15 "yes–no–I don't know" questions.

The comprehension questions are designed to assess readers' performance on four central components of the comprehension processes:

1. *Knowledge access:* Accessing relevant prior knowledge from long-term memory (e.g., "An orange has a peel.")
2. *Text memory:* Recalling from memory new information presented in the text (e.g., "Maria likes to eat fruit.")
3. *Inferences:* Making novel inferences based on information provided in the text but without prior knowledge (e.g., "A nura is smaller than an orange.")
4. *Integration:* Integrating accessed prior knowledge with new text information (e.g., "You peel a nura to eat it.")

Participants are encouraged to read the text carefully at their own pace and to answer the comprehension questions without having the text in front of them. The task is preceded by a practice text and some comprehension questions across each category. The scores are based on the number of correct answers in the four categories of questions related to the basic processes underlying reading comprehension. For the Spanish version of the DARC, the coefficient of reliability for the total score was .87.

WM

A Spanish version of the RST (Daneman & Carpenter, 1980; Spanish version: Orjales, García-Madruga, & Elosúa, 2010) for primary school students was used. In this task, participants are asked to read a series of sentences presented on a computer screen aloud and then recall the last word of each sentence in the correct order. The sentences are very simple and easy to read, using familiar words. The task includes diverse levels in which the number of sentences progressively increases from two to six. There are three series of sentences in each level. The scoring procedure was developed by Elosúa, García-Madruga, Gutiérrez, Luque, and Gárate (1997) for the RST. This procedure scores the number of words that participants are able to remember with minimum consistent performance. In each of the three series at each level, a participant's performance can be one of the following:

- Correct (accurate words, correct order)
- Half correct (accurate words, incorrect order)
- Incorrect

The minimum consistent performance at each level is reached when a participant performs at least half of the maximum—that is, either three series of words half correct or one series of words correct, one half correct, and one incorrect. Every performance better than the minimum consistent performance, at the same or higher levels, was scored by the addition of tenths of a point. In the same level, each supplementary correct response would add 0.2 point and each supplementary half-correct response 0.1 point. For instance, minimum consistent performance at the third level is 3, and maximum performance at the third level is 3.3; if a participant remembers only two of the three series of three words (level 3) accurately and in the correct order, his or her scoring would be 3.1. At a higher level, a supplementary correct response would add five-tenths of a point and a supplementary half-correct response four-tenths. For example, if the previous participant also remembers a series of four words (level 4) but in the incorrect order, his scoring would be 3.1 + 0.4 = 3.5.

Intelligence

We used the Matrices subtest of the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 2000). This test evaluates nonverbal fluid intelligence: It assesses a student's ability to solve new problems by perceiving relationships and completing abstract analogies. Because items contain pictures and abstract designs rather than words, you can assess nonverbal ability even when language skills are limited. Full-color items appeal to students, particularly those who are reluctant to be tested. For the Spanish version of the Matrices subtest, for the 8-year-old students, the reliability coefficient was .80.

Training Program

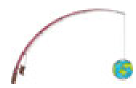
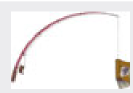
Students engaged in training on a variety of reading tasks that were especially designed to tap into the four executive functions (i.e., focusing, switching, connecting with long-term knowledge and updating mental representations, the inhibition of irrelevant information) for approximately 50 minutes a day for 12 days over a four-week period.

The battery of tasks included in the training enabled us to systematically vary demands on the executive abilities required to perform these successfully in different proportions. The focusing function is present in all of the tasks. This is because it demands students to focus their attention on specific and relevant information to resolve the task. The switching function is particularly required on the tasks in which readers have to shift back and forth between diverse pieces of information or when the task includes diverse subtasks. Connecting with long-term knowledge is particularly necessary when performing tasks that require combining information from the task with information from long-term memory. The function of updating mental representations is particularly present in those tasks that require monitoring and coding incoming information relevant to the task at hand and then appropriately revising the items held in WM by replacing old, no longer relevant information with newer, more relevant information (Morris & Jones, 1990). Finally, the inhibition of irrelevant information concerns tasks in which students need to inhibit or override the tendency to produce a more dominant or automatic response. The tasks used to tap into each executive function are presented in Table 1. Their content and what the students were expected to do in each task are described below.

The training was carried out by two researchers in the classroom during an ordinary scholastic period. At the beginning of each session, students received a workbook in which they had to fill out the solutions to the tasks completed during the session. In the first session, one of the researchers explained in a detailed and direct way the component processes as well as the outcome of reading comprehension. For this purpose, the instructor used a sentence by Miller (1977), "The Smiths saw the Rocky Mountains while they were flying to California" (p. 400), that was adapted to the cultural features of Spanish students living in Madrid: "*Laura vió la Sierra de Navacerrada mientras volaba a Barcelona*."

Participants understood and agreed that comprehending this sentence implies the participation of diverse mental processes and the integration of text information and prior knowledge to build a representation that shows a girl, "Laura," seeing the "*Sierra de Navacerrada*" through the small window of an airplane.

TABLE 1
The Executive Processes Trained, their Icons, and the Tasks Used in Experiment 1

Executive function	Icon	Tasks tapping into each executive function
Focusing		• Analogies • Anaphora • Changing stories • Decoding instructions • Inconsistencies • Inferences • Integrating knowledge • Main idea • Sentences in order • Vignettes in order
Switching		• Analogies • Anaphora • Inconsistencies • Inferences • Integrating knowledge
Connecting with prior knowledge		• Analogies • Anaphora • Changing stories • Decoding instructions • Inferences • Main idea • Sentences in order • Vignettes in order
Semantic updating in working memory		• Analogies • Anaphora • Changing stories • Inconsistencies • Inferences • Integrating knowledge • Sentences in order
Inhibition		• Analogies • Anaphora • Changing stories • Decoding instructions • Inconsistencies • Integrating knowledge • Main idea • Sentences in order • Vignettes in order

Researchers tried to gradually increase the difficulty of the tasks and the items within each task, adopting an adaptive training perspective despite the obvious limitations of collective training in a classroom setting. Students performed different tasks each day, selected from a bank of 10 kinds of tasks: vignettes in order, sentences in order, decoding instructions, anaphora, analogies, inconsistencies, inferences, main idea, changing stories, and integrating knowledge.

In the vignettes in order and sentences in order tasks, students were asked to organize either series of vignettes or series of sentences into the correct order to create a coherent story. The decoding instructions task required them to interpret and perform complex written instructions involving the integration of a sequence of actions. To do that, they had to read the instructions presented on a screen and then either write down or draw the information received in their workbooks.

In the anaphora and verbal analogies tasks, students had to solve either syntactic and semantic anaphora or analogy problems and then store and remember the word solution in a growing series of inferential problems (for a complete presentation of the anaphora and analogies WM tasks and materials, see Gutiérrez-Martínez, García-Madruga, Carriedo, Vila, & Luzón, 2005). Students had to silently read the anaphora and analogy problems presented on a screen, recall the word solution of each anaphora or analogy problem, and then write them down in the correct order.

The inconsistencies task required students to act as a detective whose job consisted of looking for mistakes in the texts. They read texts containing an internal inconsistency (i.e., an inconsistency between two ideas expressed within the text) and an external inconsistency (i.e., information that conflicted with their prior knowledge), and their assignment consisted of detecting one inconsistency of each type within each text. When performing the inferences task, students had to read different short texts presented on a screen and answer embedded questions that either required the integration among individual sentences in the text (i.e., text-based inferences) or demanded the integration of general knowledge with information in the text (i.e., elaborative inferences).

In the main idea task, students had to either locate the more important ideas of different reading passages or select the best summary from the passage. In the changing stories task, students read different texts, including a stream of information in which the relevant facts are constantly changing. They were asked to actively keep track of the information as they read it because at several points in the story, they were requested to determine the state of different aspects of the story at that time (e.g., order of the horses in a race, state of the scoreboard during a football match).

Finally, the training program included the integrating knowledge task. This activity required students to focus and switch their attention to different units of information presented on a screen in different formats (i.e., text, video, pictures) to be able to answer several questions that required the integration of multiple sources of information.

All of the tasks consisted of several items presented in order of increasing difficulty. Each task was trained by means of four modes of instruction: (1) explicit instruction in the executive functions related to the task, (2) modeling examples, (3) guided practice, and (4) independent practice. We review each of these next.

Explicit instruction was provided by one of the researchers, who explained to the students how to perform each task as well as requested them to reflect on how one might use the different executive functions to perform the tasks effectively. To make each executive function concrete and easy to understand, icons (or symbols) were used to represent them. These icons were illustrated graphically and presented to the students throughout the training program. Concretely, focusing was illustrated as a magnifying glass, switching as two eyes looking in different directions, connecting with long-term knowledge as a fishing rod with a globe, updating mental representations as a fishing rod with a book, and inhibition of irrelevant information as a stop sign (see Table 1).

The second vehicle of instruction was modeling examples. After providing explicit instruction, the experimenters completed the first two items of each task aloud, making sure that students understood what they had to do. The third mode of instruction was guided practice, which asked students to perform some items included in the task (i.e., two or three, depending on the task) while receiving visual feedback from the experimenters about the correct answers.

Finally, students completed each task performing a set of items as independent practice. It should be noted that the first day of the program, training on each task included all four of these modes of instruction (i.e., explicit instruction, modeling, guided practice, independent practice). From the second day of each task, the students participated in the training by completing the items independently. Thus, the focus of the training was independent practice. To be sure that students remember what they had to do while performing each task, the second and subsequent days of each task started with the solution of the last item completed the previous day.

To keep students motivated throughout the program, at the end of each session, they performed the motoric instructions. To do this, they had to read some instructions presented on a screen and then execute funny postures and movements with their body. Additionally, at the end of each week of training, students were awarded with a diploma and a small gift—some stickers and a comic book (*Gormiti*).

Results

The results of the four variables in the control and experimental groups can be seen in Table 2. All of the statistical comparisons were two tailed unless otherwise stated. In pretesting, there were no reliable differences between the two groups for any of the variables: reading comprehension, WM, and intelligence (Mann–Whitney's tests: DARC: $U = 116, p = .87$; RST: $U = 114, p = .81$; KBIT: $U = 81.5, p = .13$).

In the control group, the light gains obtained in posttesting with the DARC, RST, and KBIT were not

TABLE 2
Means (and standard deviations) of the Three Measures in the Pretest and Posttest and the Students' Increases for the Control and Experimental Groups in Experiment 1

	Control group (N = 16)			Experimental group (N = 15)		
Measure	Pretest	Posttest	Increase	Pretest	Posttest	Increase
Diagnostic Assessment of Reading Comprehension	31.69 (5.78)	31.81 (5.81)	0.13 (3.56)	31.27 (4.83)	34.27 (5.08)	3.00* (4.47)
Matrices subtest of the Kaufman Brief Intelligence Test	30.56 (4.13)	31.06 (3.53)	0.50 (4.70)	27.73 (4.18)	31.20 (3.55)	3.47** (4.10)
Reading span test	2.71 (0.53)	2.83 (0.62)	0.13 (0.59)	2.64 (0.49)	2.81 (0.48)	0.17 (0.56)

* $p < .05$. ** $p < .01$.

reliable (Wilcoxon tests: $z = -.42$, $p = .67$; $z = -.82$, $p = .41$; and $z = -.69$, $p = .49$, respectively). On the contrary, in the experimental group, there were reliable gains after intervention in the posttests for reading comprehension (DARC: $z = -2.179$, $p = .029$, Cohen's $d = .67$). There were also gains in the measure of intelligence (KBIT: $z = -2.642$, $p < .01$, Cohen's $d = .86$), but none were found for WM (RST: $z = -1.219$, $p = .22$, Cohen's $d = .30$). Moreover, the gain in reading comprehension was reliably higher for the experimental group than for the control group (DARC: $U = 66$, $p = .032$, Cohen's $d = .72$), and the gain in intelligence was also higher for the experimental group than for the control group, but it did not reach the significance level (KBIT: $U = 78.5$, $p = .10$, Cohen's $d = .68$). The gain found in favor of the experimental group for WM was light and not reliable ($U = 116.5$, $p = .89$, Cohen's $d = .07$).

As predicted, there was a clear pattern of positive correlations in pretest among reading comprehension, WM, and intelligence. Reading comprehension reliably correlated with WM ($r = .34$, $p < .05$, one tailed); however, the positive correlation with intelligence did not reach the significance level ($r = .21$). The correlation between WM and intelligence, although positive, also did not reach the significance level ($r = .21$).

Discussion

In this experiment, we applied to intervention the cognitive theory that points to the role of executive processes in reading comprehension. The results suggest that it is possible to improve text comprehension by training young students on executive processes during the reading process. There was a reliably higher pretest to posttest gain in the experimental group, compared with that of the control group, for reading comprehension, and this effect was between medium and large. This gain was yielded because of the training program. Thus, our results demonstrate that it is possible to develop interventions to promote reading comprehension by boosting the central executive functions during the process of reading comprehension.

There is another relevant result, as found in other studies (see Jaeggi et al., 2008; Klingberg et al., 2002): The intervention on executive functions may also improve fluid intelligence measures, in particular on the visual matrices scale of the KBIT (Rueda, Posner, & Rothbart, 2005; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005). The gain in this measure of intelligence was higher, although not reliably so, for the experimental group than for the control group, and this effect was almost as large as that found for reading comprehension.

In the posttesting, we did not find any gain in RST for the experimental group. The lack of improvement in this WM measure might be due to the particular characteristics of the RST. As we noted in the introduction, the RST is a WM capacity measure that loads mainly on storage and verbal components. We need some better measures of WM's executive processes, which demand more attentional resources, to test an improvement in these processes.

We also confirmed our predictions about positive correlations among the studied variables. Reading comprehension is reliably related to WM and not significantly with intelligence. The positive correlation between WM and intelligence also did not reach the significance level. This low correlation can be explained by the kind of measures used for WM and intelligence. As we just mentioned, the RST is a central executive measure, but it loads mainly on storage and verbal components of WM. These components are not particularly related to fluid intelligence as measured by the KBIT matrices.

This experiment has some obvious limitations, particularly the reduced number of participants, the lack of a more complete measure of WM, and the large time lapse between pre- and posttest measures. In experiment 2, we intended to overcome these limitations by using a larger experimental group, three diverse measures of WM and the central executive, and a shorter time period (less than three months) for the entire pretest–intervention–posttest period. The main aim of experiment 2 was to analyze in a detailed way the effect and the process of training on participants according to their prior reading comprehension abilities. It also attempted to relate pre- and posttest measures with some measures of the training process itself. Moreover, from the experience of applying the training program in experiment 1, some modifications in the training process were also included in the new experiment.

Experiment 2

There were three objectives of this experiment. First, we intended to confirm the efficacy of the training program and to analyze in a deeper way the tasks used in the program and their relations with the efficacy of the intervention. For this purpose, during the training process, we scored and analyzed participants' performance in each of the tasks used to improve reading comprehension. Second, we wanted to check the possible diverse effect of training on participants according to their prior abilities on reading comprehension. To achieve this goal, we divided the global group of participants into two groups of low and high reading comprehension abilities, according to their DARC scores in pretesting. Third, we were also interested in the particular effect of training on each of the three main components of reading comprehension measured by the DARC, namely, memory, inferences, and integration.

As mentioned, apart from using the same measure of intelligence (visual matrices of the KBIT), we used a lightly modified version of the DARC and three new measures of WM and executive processes: a verbal analogy span test, a semantic updating span test, and a visuospatial selective span test. These three measures load more control executive processes than does the RST used in experiment 1. As in experiment 1, our objective was to assess the possible transfer of training to measures of intelligence and WM's executive processes.

We had three hypotheses:

1. There will be a significant increase after training in the posttest measure of reading comprehension (DARC) and its component measures. We also predicted a pretest to posttest increase in the KBIT's Matrices subtest of intelligence and in the new measures of WM and executive processes.
2. Given that training was collectively carried out in the classroom, training is particularly adapted to the low reading comprehension abilities group. Therefore, the increase in reading comprehension will be higher in the low reading comprehension group than in the high reading comprehension group.
3. There will be positive correlations in the pretest among reading comprehension, WM, and intelligence. Likewise, we predicted positive correlations between pretest measures in reading comprehension, WM, and intelligence and participants' performance on the training task.

Method

Design and Participants

We used an intervention design with pretraining and posttraining measures with an experimental group. Forty-six third-grade students from the same school as in experiment 1, ages 8–9 years, participated in the experiment. The final number of participants who completed at least nine of the 10 training sessions was 40 ($M = 8.61$, $SD = 0.28$). To evaluate our second hypothesis, we divided participants into two groups according to the median score obtained in the DARC pretest: low ($n = 21$, mean age = 8.49 years, mean DARC score in pretest = 18.62) and high ($n = 19$, mean age = 8.70 years, mean DARC score in pretest = 30.89) groups in reading comprehension abilities.

Procedure

The students were trained for 10 days in their classroom over a four-week period on a number of reading comprehension tasks directly related to one or more of the central executive functions (see a description of the tasks in the Experiment 1 section). During the training process, as in experiment 1, all participants received a workbook that included the diverse exercises to be performed in each session. These workbooks were collected by experimenters at the end of each session. Participants were assessed on three measures of WM (analogy, semantic updating, and visuospatial tests) and on measures of intelligence (matrices of the KBIT) and reading comprehension (DARC) two weeks before and after training.

Pretest and Posttest Measures

Reading Comprehension

To assess reading comprehension, we used the Spanish version (EDICOLE) of the DARC, as in the previous experiment. However, we introduced a light modification to the EDICOLE. In particular, we asked for three different relations between real and unreal entities in all of the texts so that in this experiment, the test had the same number of questions per category (i.e., prior knowledge, text memory, inference, integration) across texts, even though the number of total comprehension questions per text did not vary with respect to the first version. Because almost 100% of participants performed correctly on the prior knowledge items, to calculate the overall DARC score, we did not consider the knowledge scores in experiment 2. For the new version of the DARC, the coefficient of reliability for the total score was .84.

WM and Executive Processes

A new analogy test of WM for primary school students (Orjales & García-Madruga, 2010) was used to assess students' WM capacity. In this task, participants are asked to read aloud and solve a series of verbal analogies and then recall the word solution of each analogy in order. The verbal analogies are very simple and easy to solve, for example: "Teacher is to school as doctor is to: a) medicine; b) hospital."

The structure of the task is very similar to the RST. However, in this case, instead of only reading aloud and automatically selecting the last word of each sentence, participants have to solve a verbal analogy inference and store and remember the correct word solution. The task includes diverse levels in which the number of verbal analogies to be resolved by participants progressively increases from two to five. There were three series of verbal analogies in each level. The scoring procedure was different from that of experiment 1. We gave 1 point for each word remembered in a correct series in which participants remembered all the words in correct order. When participants remembered all the words of a series but changed the order of some words, the words changed in order were scored as 0.5 point. The recall of correct words when participants did not remember all the words of a series was not considered.

Semantic Updating Test

Based on the work of Palladino et al. (2001), we developed a semantic updating test for primary school students, in which the updating process relies on a semantic criterion to make the task as similar as possible to the updating process involved in reading comprehension. This task assesses the recall of a variable number of items following a specific semantic criterion in a list of words. Participants are presented with nine lists that include eight concrete and highly familiar words that refer to objects, vegetables, or animals measurable by size. Students are required to select and remember a limited and predefined number of the biggest elements that were named in the word list while suppressing the rest of the elements.

The nine lists are divided into three levels of increasing trials, varying by the number of relevant elements to be recalled (i.e., two, three, or four words). The lists are presented in a fixed order. The words in a list (e.g., *elephant*, *pea*, *lightbulb*, *phone*, *glasses*, *train*, *tooth*, *pencil*) are presented on a computer screen, at an approximate rate of two seconds per word, while they are also named aloud by the experimenter. The end of the list is signaled graphically on the screen, and the participant is immediately required to report on the two, three, or four words in the task referring to the biggest objects or animals following the order in which they were presented in the list.

The instructions emphasize that the participant will be presented with lists that include eight nouns referring to animals or objects, the size of which has to be considered to select the predefined number of the biggest elements at each moment. Participants are not informed about the range of positions within the list where the target items are going to appear, such that they must pay attention to all positions. To carry out the task successfully, participants have to change the content of memory by updating irrelevant old items with relevant incoming items (the biggest element). The task is preceded by three practice lists of two-word strings that participants must remember. Returning to the example given above, the participants would have to recall *elephant* and *train*. The scoring procedure was the same as in the previous analogy WM test.

Visuospatial WM Test

A new test of visuospatial WM was used for assessing students' visuospatial WM capacity and the executive processes related to the control of the dual task. The visuospatial test is a Spanish adaptation of the visual span task developed by Cornoldi et al. (2001). The test consists of a series of locations of several black dots presented in 4×4 (16 cells) white matrices in which one of the rows and one of the columns of each matrix randomly appears colored in gray. Positions of dots are randomly distributed in the matrix's cells and held visible for two seconds on the screen. When the last dot of each series is displayed, a bell rings to inform participants of the end of each series. The task has three levels of difficulty. Each level consists of three series of increased number of trials (i.e., positions of dots in the matrix), ranging from two to four trials.

Participants are asked to do two tasks simultaneously: (1) press the spacebar when the black dot appears in a gray cell in the matrix (do not press it when the dot appears in a white cell); and (2) at the end of each series, remember the positions of the last dots of each series in order of appearance and identify them on a new blank matrix. Thus, the positions of dots that participants have to remember are only those that appear in the matrix when the bell rings: two dots in level 2, three in level 3, and four in level 4. The scoring procedure was the same as in the analogy and semantic updating tests, except what counts is remembering the dots in their positions, not words. The remembering of correctly placed dots in an incorrect order was also scored as 0.5 point.

Intelligence

We used the same Matrices subtest of the KBIT as in experiment 1.

Training Program

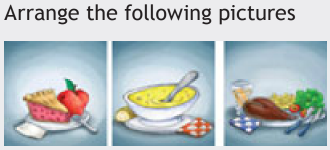
The training program described in experiment 1 was reviewed to design a more adjusted and shorter program (i.e., 10 sessions instead of 12). The four criteria used to improve the intervention were the following:

1. To eliminate those tasks and items with less satisfactory results by considering the difficulty and homogeneity indexes obtained for task training scores
2. To reformulate the statements of some items while taking into account the misunderstandings observed in some of the students who participated in experiment 1
3. To adjust the level of difficulty of each task to avoid ceiling and floor effects that were detected in some of the items
4. To increase the number of items in guided and independent practice for those tasks that showed better statistical results

Table 3 shows the changes resulting from this revision.

The new intervention program consisted of ten 50-minute sessions across four weeks. As shown in Table 3, the main idea and analogies tasks were not present in this new version of the intervention, thus the number of tasks was now eight. The order of items and

TABLE 3
Training Tasks, Examples, Variables Manipulated for increasing the Difficulty, Sessions in Which Each Task Was Performed in Experiment 2, and Number of Items in Experiments 1 and 2

Task	Example of task item	Difficulty	Sessions	Number of items: Exp. 1	Number of items: Exp. 2
Anaphora working memory	*Robert painted it white before the summer arrived.* *– roof* *– façade*	• Number of words to be remembered	4, 5	14	14
Detecting textual inconsistencies	Internal: *Laura used eyeglasses to read.... Laura's eyesight was excellent.* External: *Elena was flying in the depths of the lake when he decided to go back.*	• *Internal*: Distance between sentences • *External*: Salience of the inconsistency	5–7	16	30
Decoding written instructions	*Write your name and two surnames. Then, draw a circle around the last letter of your name and the first letter of your last surname. Do it without lifting your pencil.*	• Number of actions to be performed	2–10	45	48
Following changing stories	*In what order were the horses at the end of the race?*	• Number of units of information to be followed	8, 9	12	18
Integrating information from different formats	After watching the video and reading the text, ask the following question: *What type of solar eclipse is presented in that picture?*	• Number of units of information to be integrated across sources	8, 9	15	15
Making inferences	After the student reads the text, ask the following questions: *Why did they put the sparrow near the fireplace?*	• *Text-based*: Distance between sentences • *Elaborative*: Memory load	6, 7	16	30
Sentences in order	Arrange the following sentences: *Maria looks for her place.* *Maria buys the ticket.* *The movie has started.* *Maria waits in the line.*	• Number of sentences	3, 4	12	26
Vignettes in order	Arrange the following pictures 	• Number of frames	1, 2	19	50

tasks was rearranged to adjust the increasing level of difficulty of items and tasks to students' performance.

Participants were trained on each task by using the same mode of instruction and following the same procedure as described in the previous training program. The motoric instructions to be performed throughout the program, as well as the diploma and a small gift awarded at the end of each week, were also maintained to ensure students' motivation in this new training program. In the last session of training, we illustrated and had students reflect on the utility of the four basic executive processes for diverse daily intellectual activities; likewise, we insisted on the idea that the repeated practice of the four basic processes were developed so students could become "mental athletes." In this final session, a diploma was presented to each of the students.

Results

Table 4 shows the effect of training on reading comprehension measures. As in experiment 1, all of the statistical comparisons were two tailed unless otherwise stated. There were reliable gains after training on the three reading comprehension measures of memory, inferences, and integration, as well as on the overall measure of the DARC. The effects of training on the diverse component measures of reading comprehension were around medium size; larger and more significant effects were found for integration and inferences, and smaller and less significant effects were found for memory. The effect size for the overall DARC was large and greater than in experiment 1.

The effect of training on the rest of the measures can be observed in Table 5. As in experiment 1, there was a reliable increase after training on the visual matrices of the KBIT, and the effect was medium to large. The gains of the semantic updating and visuospatial tests were also reliable, and the effects were from medium to large. The gain obtained in the analogy WM test after training did not reach the significance level, and its effect was small.

Regarding the second hypothesis, Figure 1 shows the scores of low and high reading comprehension groups on the DARC and KBIT. As can be observed, there were clear differences between both groups in both variables, although in opposite directions. The main increase in reading comprehension was obtained by the low reading comprehension group, whereas the main increase in intelligence was obtained by the high reading comprehension group. The gain in reading comprehension of the low group was greater than the gain of the high group, and the effect was very large (Mann–Whitney's test: $U = 73.5$, $p < .001$, Cohen's $d = 1.34$); on the contrary, the gain in intelligence was greater for the high reading comprehension group than for the low reading comprehension group, and the effect was medium to large ($U = 123.5$, $p < .05$, Cohen's $d = 0.69$).

TABLE 4
Means (and standard deviations) of the Measures of Reading Comprehension in the Pretest and Posttest and the Students' Increases (with Cohen's *d*) in Experiment 2 (*N* = 40)

	DARC memory	DARC inferences	DARC integration	DARC
Pretest	9.57 (2.02)	6.18 (3.09)	8.70 (3.92)	24.45 (7.78)
Posttest	10.23 (1.69)	7.53 (2.54)	11.00 (3.07)	28.76 (6.27)
Increase	0.66* (2.05) *d* = 0.33	1.35** (2.24) *d* = 0.62	2.30** (3.63) *d* = 0.65	4.31** (5.65) *d* = 0.79

Note. DARC = Diagnostic Assessment of Reading Comprehension.
* $p = .05$. ** $p < .001$.

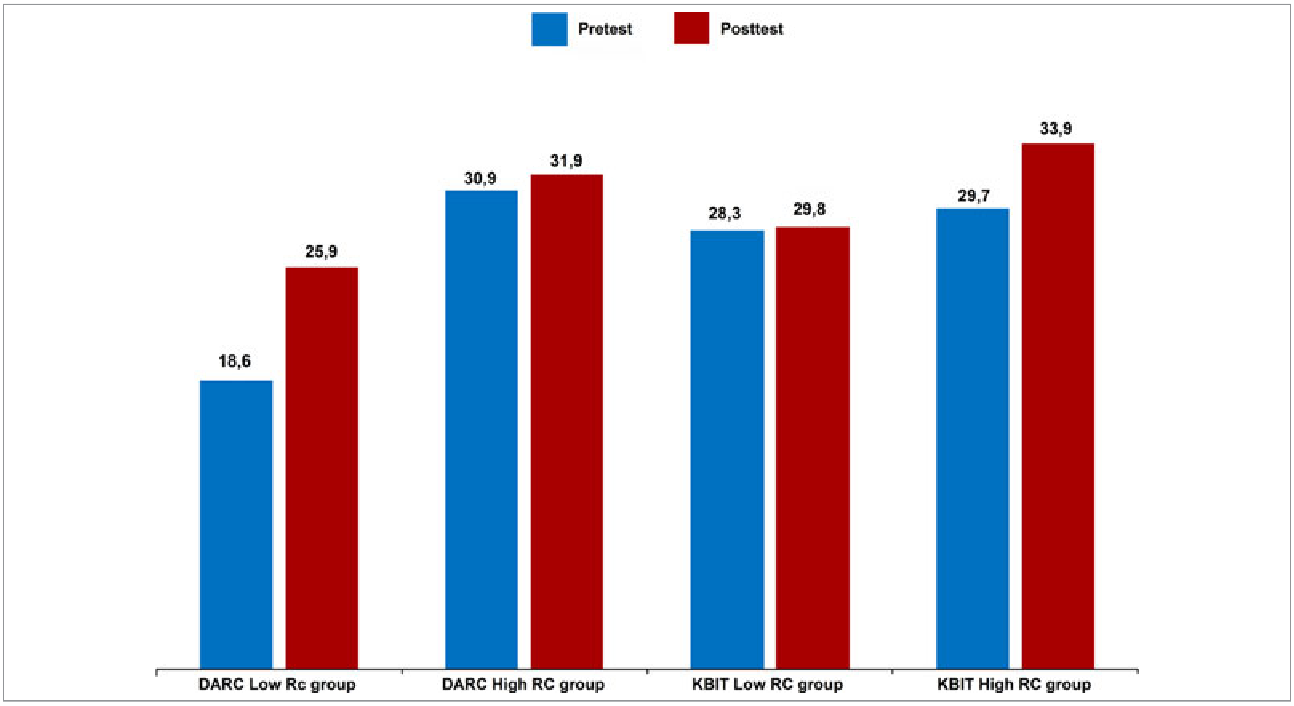
TABLE 5
Means (and standard deviations) of the Measures in Pretest and Posttest and the Students' Increases (with Cohen's *d*) in Experiment 2 (*N* = 40)

	Kaufman Brief Intelligence Test	Analogies working memory test	Semantic updating test	Visuospatial span test
Pretest	28.98 (3.90)	5.55 (3.56)	8.60 (5.00)	13.58 (6.61)
Posttest	31.73 (4.85)	6.58 (4.49)	12.70 (5.89)	18.94 (5.66)
Increase	2.75** (4.24) *d* = 0.66	1.03† (4.18) *d* = 0.25	4.10** (6.66) *d* = 0.62	5.36** (6.98) *d* = 0.77

** $p < .001$. † $p = .10$.

FIGURE 1

Mean Scores on the Diagnostic Assessment of Reading Comprehension and the Kaufman Brief Intelligence Test for Low and High Reading Comprehension Groups in Pretest and Posttest in Experiment 2



To test in a stricter way this new finding, that there was a greater gain in intelligence by the high reading comprehension group, we calculated the posttest–pretest/pretest scores in both groups. The means (and standard deviations) were: 0.15 (0.15) for the high group and 0.06 (0.16) for the low group. The Mann–Whitney test showed the difference once again, although now it was marginally reliable (U = 128, p = .053, Cohen's d = 0.58). There were no significant differences between the two groups in the gains for the three WM measures: analogies, semantic updating, and visuospatial span.

The correlations among reading comprehension, intelligence, and a composite measure of WM are shown in Table 6. The composite measure of WM was the mean of z-scores of the three measures: analogy, semantic updating, and visuospatial tests. As predicted, reading comprehension reliably correlated with WM (r = .29, p < .05, one tailed) and intelligence (r = .39, p < .01, one tailed). However, the correlation between WM and intelligence, although positive, did not reach the significance level (r = .15).

Table 6 also shows the correlations of the three components and the overall measure of the DARC with the three WM measures, the composite measure of WM, and the KBIT measure of intelligence. As can be observed, the correlations among the three component

TABLE 6

Pearson Correlations in the Pretest of the Four Measures of Reading Comprehension (memory, inferences, integration, and overall) with Intelligence (KBIT) and the Four Measures of Working Memory (analogy span, semantic updating span, visuospatial selective span, and composite score) in Experiment 2 (*N* = 40)

	KBIT	Analogy test	Semantic updating test	Visuospatial selective test	Composite working memory
DARC memory	.29*	.44**	.05	−.06	.22
DARC inferences	.31**	.27*	.34*	.01	.31**
DARC integration	.29**	.05	.28*	.11	.22
DARC	.39**	.25	.28*	.05	.29*

Note. DARC = Diagnostic Assessment of Reading Comprehension. KBIT = Kaufman Brief Intelligence Test.
* p < .05, ** p < .01, one tailed.

TABLE 7
Percentages of Correct Responses (means and standard deviations) in the Tasks of the Training Program, and Pearson Correlations with Reading Comprehension (DARC), Intelligence (KBIT), and Semantic Updating in Pretest in Experiment 2 (*N* = 40)

	Mean	*SD*	DARC	Semantic updating test	KBIT
Anaphora	57	25.80	.53**	.46**	.22
Changing stories	72	23.12	.31*	.33*	.20
Decoding instructions	75	13.74	.46**	.33*	.35*
Inconsistencies	80	11.69	.28*	.42**	.04
Inferences	74	8.54	.41**	.32*	.22
Integrating knowledge	84	20.51	.28*	.21	.08
Sentences in order	85	13.53	.36*	.41*	.29*
Vignettes in order	54	15.24	.16	.34*	.06
Overall training tasks	73	10.77	.55**	.54**	.28*

Note. DARC = Diagnostic Assessment of Reading Comprehension. KBIT = Kaufman Brief Intelligence Test.
* $p < .05$, ** $p < .01$, one tailed.

measures of the DARC and WM measures confirmed our hypothesis, except for the unsurprising case of the visuospatial measure. The analogy test significantly correlated with memory and inferences, and the semantic updating test correlated significantly with inferences and integration. The KBIT reliably correlated with the three component measures of the DARC.

The results obtained by participants in the diverse training tasks during the process of intervention, as well as the correlations with reading comprehension, semantic updating, and intelligence, can be observed in Table 7. Arranging vignettes in order was the easiest task, whereas arranging sentences in order was the most difficult. An analysis of the intercorrelations among the diverse training tasks showed that vignettes in order did not correlate significantly with any other task. The intercorrelations among the other seven tasks were always positive, and all were significant except for two cases. Likewise, all the training tasks, except for the vignettes in order task, correlated significantly with reading comprehension in the pretest.

Semantic updating in the pretest also clearly correlated with the diverse training tasks, and we can observe that its correlation with the overall training task is almost as high as that obtained by reading comprehension. The correlation between the KBIT in the pretest

TABLE 8
Pretest's and Posttest's Pearson Correlations Between the Overall Training Tasks Measure and the Four Measures of Reading Comprehension (memory, inferences, integration, and overall), Intelligence (KBIT), and the Composite Measure of Working Memory in Experiment 2 (*N* = 40)

	Overall training tasks	
	Pretest	Posttest
DARC memory	.24	.18
DARC inferences	.56**	.49**
DARC integration	.52**	.35*
DARC	.55**	.42**
KBIT	.28*	.36*
Composite working memory	.29*	.38**

Note. DARC = Diagnostic Assessment of Reading Comprehension. KBIT = Kaufman Brief Intelligence Test.
* $p < .05$, ** $p < .01$, one tailed.

and the overall training tasks was also reliably positive; the correlations with the diverse component measures of training were also positive, although in most cases, they did not reach the significance level.

To assess the predictive capacity of these three pretest tasks on students' performance in the training tasks overall, we carried out a multiple regression analysis following the stepwise method. The results showed that reading comprehension and semantic updating explained 43% of the variance of the performance in training tasks: $F(2, 37) = 15{,}722$, $p < .0001$; both variables, the DARC and semantic updating, were significant: $\beta = 0.43$, $p < .0002$; and $\beta = 0.42$, $p < .0002$, respectively.

Finally, the correlations between participants' overall performance on the training tasks and the four DARC measures, intelligence, and composite WM on the pretest and posttest can be seen in Table 8. There was a clear pattern of positive correlations among the diverse measures, with a range between .18 and .56. As predicted, the correlations between the diverse reading comprehension measures and the overall training tasks score were positive and significant both in pretest and posttest, except for the memory measure. Likewise, the correlation between the composite WM and intelligence and the overall training tasks score were positive and significant.

Discussion

The results clearly confirmed the first hypothesis showing reliable increases after training across the three components of the reading comprehension text: memory, and particularly in inference and integration.

It is unsurprising that a training program based on the executive processes involved in reading comprehension yields greater benefits on inference and integration than on memory (the size effects are almost doubled). In comparison with memory, the inference and integration components of reading comprehension are more difficult: They require an extra mental operation, and therefore executive control is more involved. The changes introduced in the training program seem to have improved it because the effect size is larger than that obtained in experiment 1, and with two fewer sessions.

Results also confirmed the finding of experiment 1 regarding the increase in participants' fluid intelligence after training the executive processes involved in reading comprehension. Moreover, in this case, the transference is extended to WM measures: We also found a reliable increase after training in semantic updating and visuospatial WM tasks and no significant gain on the analogy test. The use of the new tasks has allowed us to find reliable gains in WM's executive processes, although we are unable to find an easy explanation for the lack of reliability in gains in the analogy task. These kind of transfer effects, whereby WM training improves intelligence, is interpreted by some authors as consistent with the evidence of a probable common or overlapping frontoparietal cortical network involved in intelligence, WM, and executive processes (see Duncan, 2010; Klingberg, 2010).

The second hypothesis regarding the differential efficacy of the training program according to participants' prior reading comprehension abilities has also been confirmed. The low reading comprehension group reached a very clear and reliably greater gain after training than the high reading comprehension group did. Because our training program was particularly adapted to the low reading comprehension group, this result provides further evidence in favor of adaptive training. As for the differential performance of low and high reading comprehension groups, we found a rather new result that, if confirmed in new studies, may be particularly interesting: Participants with high reading comprehension abilities seem to have used our training program to improve their fluid intelligence rather than their reading comprehension abilities.

To shed some light on this result, we also divided the whole group of participants by the median in KBIT pretest scores and compared the gains of the resulting high and low pretest intelligence groups. The increase was reliably higher for the low intelligence group than for the high intelligence group: $M = 4.48$, $SD = 4.27$; and $M = 0.84$, $SD = 3.35$, respectively; Mann–Whitney's test: $U = 109.5$, $p < .02$, Cohen's $d = 0.96$. Therefore, the gain after training in the KBIT seems to be greater for students with low intelligence and high reading comprehension abilities. These results suggest an interesting role of reading comprehension on improving intelligence.

Our third hypothesis was confirmed on the whole. The correlations among diverse measures in pretesting were in the predicted direction, although, as in experiment 1, WM and fluid intelligence were not reliably correlated. As we noted earlier, this result seems to be related with the kind of measures used for intelligence and WM. This explanation is consistent with positive significant correlations found in the posttest between fluid intelligence (KBIT matrices) and visuospatial WM ($r = .28$, $p < .05$, one tailed). Another interesting result is the lack of correlation in the pretest between reading comprehension and visuospatial WM (see Table 6). This result is relatively common (see, e.g., Seigneuric et al., 2000), although there are also some studies that show a relationship between reading comprehension and visuospatial WM (e.g., Goff, Pratt, & Ong, 2005). Again, a possible explanation of these contradictory results relies on the differences among the diverse measures used for the variables.

The results obtained by students in the training tasks provide global support to the training program: The overall level of difficulty of diverse tasks is adequate, and there is a good consistency among the tasks, except for the vignettes in order task. This is unsurprising because this task, although similar to other tasks, is not exactly a reading comprehension task. To adequately solve the vignettes in order task, one only has to understand each picture and apply one's long-term memory script knowledge about social situations. In spite of its peculiarities, use of the vignettes in order task is, in our opinion, recommended because it is quite attractive to students: It has a clear motivational value within the training program, particularly in the first two sessions in which students have to perform it.

The most interesting result regarding the training tasks is the pattern of positive correlations found with the pretest reading comprehension and semantic updating measures. Given that the overall training score is an online measure of the training process, the medium–high correlation obtained with reading comprehension confirms that the training tasks are indeed acting on reading comprehension, something that provides consistency to our training procedure. The other result that deserves to be mentioned is the positive and reliable correlations found between training tasks and the semantic updating task. The correlation between participants' overall performance on the training task and the semantic updating task is over .50 and of the same magnitude as the correlation found with reading comprehension (see Table 7).

Moreover, the results of regression analysis confirm that students' performance on the training tasks strongly depends not only on their prior reading comprehension abilities but also on their semantic updating capacity. These results provide new evidence in favor of the crucial role that updating has in reading comprehension (see Carretti et al., 2005; Palladino et al., 2001) and illustrate in particular its relationship with improvements in reading comprehension—that is, with semantic learning.

General Discussion

Klingberg (2010) distinguishes between two types of WM training. The first he calls implicit because it "is based only on repetition, feedback and often gradual adjustment of the difficulty" (pp. 317–318). The second is an explicit type based on teaching metacognitive strategies to improve performance. The training used in this work shares features of both types: We developed a training procedure based on repetition, feedback, and the gradual adjustment of difficulty. We did not explicitly train on any particular strategy, but our training explicitly demands from students their active and conscious engagement throughout all of the training process, from the first to the final session.

In fact, the main focus of our training procedure was not to train reading comprehension itself but to train WM's executive processes—that is, the conscious control of cognitive processes involved in reading comprehension. This is a key difference between the main previous approach in the field of reading intervention programs and the approach presented herein (see Elosúa, García-Madruga, Vila, Gómez-Veiga & Gil, in press). Obviously, as a final outcome, our proposal of using repetitive practice was intended to achieve some kind of automated behavior but always under the control and monitoring of executive processes. This is exactly the objective of using the "mental athletes" metaphor to characterize our kind of perspective: the conscious use of repetitive training exercises to improve students' reading comprehension abilities.

All of the training tasks used in this work, even the WM tasks (anaphora and analogy in experiment 1 and anaphora in experiment 2), required that students understand written text: For this reason, they are also reading comprehension tasks. However, our training was not a reading comprehension program aimed to instruct readers on particular skills or strategies embodied in reading comprehension. From the first to the last session, our training sought an improvement in students' mental activation so they might apply the four WM executive processes previously mentioned to reading itself. Likewise, particular training tasks were not selected and arranged for their relevance to directly improving reading comprehension but for the involvement of executive processes—in other words, because they require increasingly higher attentional control resources and can hence improve students' use of executive processes during reading.

If our aim is to train readers on using WM executive processes, we cannot train them in the abstract. Moreover, if we want to improve reading comprehension, it seems appropriate to use some form of a reading comprehension task. It is difficult, therefore, to separate the specific differential weight of WM's executive processes training with that of reading comprehension practice to properly explain the improvement of reading comprehension found in this study.

The results of our two experiments provide support to the training perspective that WM's executive processes facilitate reading comprehension. In experiment 1, we found a clear gain in reading comprehension in the experimental group, reliably higher than that of the control group. In experiment 2, we confirmed the efficacy of a simplified and adjusted version of the training program with a broader experimental group. The results confirmed that the gain was mainly on the inference and integration components of the reading comprehension test, the components that require more active reading comprehension (i.e., the explicit application of executive processes to reading). Likewise, as predicted, we found that the gain was greater for those students with lower pretested reading comprehension abilities.

The rationale of this prediction was that our attempt to adapt an item's difficulty to participants was restricted as a result of the collective nature of the intervention: We were forced to focus our training on students with lower ability. Moreover, this differential gain of low versus high reading comprehension groups cannot be attributed to a kind of ceiling effect on the high group: The mean score of the high group after training (33.9) is not close enough to the maximum score (39), and only two participants reached this maximum score. In any case, the findings herein suggest the use of this training perspective as part of classroom instruction mainly to improve reading comprehension in students with poorer abilities.

Nevertheless, the finding confirming the efficacy of the training program in reading comprehension is not the only relevant result of this work. We also found evidence for a transfer effect of our training on intelligence that may be quite relevant. In both experiments, we found a reliable gain after training on our measure of fluid intelligence: KBIT matrices. As mentioned previously, this confirms other similar results found by diverse authors in students with and without ADHD (Klingberg et al., 2002; Rueda, Posner, & Rothbart, 2005; Rueda, Rothbart, et al., 2005), young adults (Jaeggi et al., 2008), and even in the elderly (Borella, Carretti, Riboldi, & De Beni, 2010), although other relevant studies have not found this transfer effect of training on intelligence (e.g., Holmes, Gathercole, & Dunning, 2009).

As we discussed, in experiment 2, the differential results of a greater gain in intelligence obtained after training by the high reading comprehension group deserves closer attention and further empirical work. In this experiment, we also found a transfer effect of training on the executive process measures, having obtained reliable gains on the semantic updating and visuospatial WM tests and nearly reliable gains on the analogy test. In any case, the lack of a control group in this experiment forces us to be prudent, although the size of the effects found tells us that these findings would probably be confirmed with a more complete design.

A new feature of our training procedure is that we obtained scores in the training process itself. That is, in each session, participants received a workbook in which they had to solve and record all the problems included in each task. In this way, at the end of the training, we had participant scores in each of the tasks used as well as an overall training task score. These training process scores are certainly not posttest measures, but neither are they pretest measures. Instead, they provide us with online information about the learning process. As we have shown, these scores allow us to check the consistency of the training procedure and potentially remove particular tasks from the empirical data.

In sum, our new training perspective, based on the improvement of WM's executive functions in reading tasks, seems to be at least as useful as the best of the training programs recently developed, in spite of its being applied collectively to all the students in the classroom. As some authors have claimed (e.g., Duncan, 2010; Jaeggi et al., 2010; Klingberg, 2010; Rueda, Posner, & Rothbart, 2005), it is unsurprising that performing repeatedly complex cognitive tasks that demand the precise, deep, and controlled understanding of increasingly difficult texts yields an improvement on tasks that demand the activation of the same or overlapping cognitive processes and brain structures. In other words, we surmise that any kind of intervention that asks participants to face new tasks that require overlapping cognitive processes, particularly ones with high cognitive demands and attentional control, would yield similar results because they also produce some changes in the activity of frontal and parietal brain cortices. Hence, the underlying assumption regarding these transfer effects is brain plasticity, a notable characteristic of a child's brain.

Our results suggest some implications for classroom teaching related to the acquisition of reading comprehension. The first idea, perhaps rather obvious, is that contrary to traditional scholastic conceptions of how to teach reading, comprehension needs to be explicitly taught, at least for some students. The diagnostic assessment of students who require the explicit teaching of reading comprehension is an important preliminary step, one in which the DARC might be a useful tool. The second idea is that this explicit teaching of reading comprehension can be based on promoting the application of WM's executive processes, as we do in our training program, in a way similar to the studies of Gaskins et al. (2007) and Meltzer et al. (2007). Likewise, the new training perspective applied in this work can be used to develop computer programs that allow an individual application adapted in a more specific way to students with specific reading comprehension difficulties or even ADHD.

Limitations

This work has some obvious limitations that we would like to highlight. Following the ideas by Melby-Lervåg and Hulme (2012) and Shipstead et al. (2012) presented in our introduction, we can pinpoint four main difficulties and problems of WM training studies. The first difficulty concerns the theoretical analysis of the processes involved in the training. Our proposal clearly identifies WM's executive processes, not WM's storage or short-term span, as the crucial component that must be trained to achieve an improvement in reading comprehension and other high cognitive abilities. However, apart from the singular function our data seem to afford the process of updating, we cannot be exactly sure what the precise role of each of the executive processes analyzed in our training work is.

The second point refers to the task used in the pretest and posttest and in training. A clear limitation is the use of single tasks in the pretest and posttest to determine whether there is any improvement in cognitive ability, such as in WM capacity, reading comprehension, or intelligence. Our work has certainly used only a task to decide that our training program was able to improve intelligence (KBIT Matrices subtest) and reading comprehension (DARC), although the DARC includes three different measures of basic components of reading comprehension: memory, inferences, and integration. Likewise, in experiment 2, we used three different tasks to measure WM: analogy, semantic updating, and visuospatial tests. Another main related limitation is the use of the same tasks in pretest and posttest as in training. Our work has avoided this important flaw by using clearly different tasks in training than was used in the pretest and posttest.

The third difficulty concerns the use of active noncontact control groups. This methodological limitation directly affects the results of our first experiment. Given the high personal involvement and motivation of the students in our experimental groups, our results have to be confirmed with new studies using a design with control groups in which students feel as involved in the experiment as do those in the experimental group. For instance, we might compare our training program with two other training groups, one based on training only WM capacity tasks and the other on training diverse reading comprehension tasks.

Finally, maintaining these kinds of improvements over time is as important as achieving them in the first place. We need to investigate these maintenance effects after a delay by means of follow-up measures on the variables. Therefore, given the limitations we pointed out, we consider that our results should be confirmed in further research.

Conclusions

Reading comprehension is a highly demanding task in which WM's executive processes play a crucial role. Our work suggests that reading comprehension can be improved by training the main WM executive processes involved in reading comprehension. The adaptive program was tested in the classroom with primary school students, and the results show that gains in reading comprehension were higher for students with low pretest abilities. We also found that students improved in measures of intelligence and executive processes. This work provides new and promising, although initial, evidence confirming the possibility of improving cognitive abilities through the adaptive training of attentional control processes involved in the execution of highly demanding cognitive tasks.

NOTES

This research was conducted with financial support of the research projects DGICYT PSI2008-00754 and Consolider-Ingenio 2010 (CSD2008-00048), both from the Ministry of Science and Innovation of Spain. We thank the students who voluntarily participated in this research as well as the Spanish state school "Emilio Casado" of Alcobendas (Comunidad de Madrid). We also thank those who discussed parts of this research with us: Alan Baddeley, Erika Borella, Martin Buschkuehl, Barbara Carretti, Cesare Cornoldi, Rossana De Beni, Susan Gathercole, Graham Hitch, Susanne Jaeggi, Carmen López Escribano, and Elena Pérez Hernández.

REFERENCES

Ackerman, P.L., Beier, M.E., & Boyle, M.O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*(1), 30–60. doi:10.1037/0033-2909.131.1.30

August, D., Francis, D.J., Hsu, H.A., & Snow, C.E. (2006). Assessing reading comprehension in bilinguals. *The Elementary School Journal*, *107*(2), 221–238. doi:10.1086/510656

Baddeley, A.D. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A.D. (2007). *Working memory, thought, and action*. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780198528012.001.0001

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic. doi:10.1016/S0079-7421(08)60452-1

Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review*, *1*(1), 3–38. doi:10.1007/BF01326548

Belacchi, C., Carretti, B., & Cornoldi, C. (2010). The role of working memory and updating in Coloured Raven Matrices performance in typically developing children. *European Journal of Cognitive Psychology*, *22*(7), 1010–1020. doi:10.1080/09541440903184617

Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging*, *25*(4), 767–778. doi:10.1037/a0020683

Britton, B.K., & Graesser, A.C. (Eds.). (1996). *Models of understanding text*. Mahwah, NJ: Erlbaum.

Cain, K. (2006). Children's reading comprehension: The role of working memory in normal and impaired development. In S.J. Pickering (Ed.), *Working memory and education* (pp. 61–91). San Diego, CA: Academic. doi:10.1016/B978-012554465-8/50005-3

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31–42. doi:10.1037/0022-0663.96.1.31

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, *19*(2), 246–251. doi:10.1016/j.lindif.2008.10.002

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of poor and good comprehenders. *Journal of Experimental Child Psychology*, *91*(1), 45–66. doi:10.1016/j.jecp.2005.01.005

Chen, T., & Li, D. (2007). The roles of working memory updating and processing speed in mediating age-related differences in fluid intelligence. *Aging, Neuropsychology, and Cognition*, *14*(6), 631–646. doi:10.1080/13825580600987660

Colom, R., Abad, F.J., Quiroga, M.A., Shih, P.C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, *36*(6), 584–606. doi:10.1016/j.intell.2008.01.002

Cornoldi, C. (2006). The contribution of cognitive psychology to the study of human intelligence. *European Journal of Cognitive Psychology*, *18*(1), 1–17. doi:10.1080/09541440500215889

Cornoldi, C., Marzocchi, G.M., Belotti, M., Caroli, M.G., De Meo, T., & Braga, C. (2001). Working memory interference control deficit in children referred by teachers for ADHD symptoms. *Child Neuropsychology*, *7*(4), 230–240. doi:10.1076/chin.7.4.230.8735

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9781139174909.006

Cowan, N. (2005). *Working memory capacity*. New York: Psychology.

Dahlin, K.I.E. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing*, *24*(4), 479–491. doi:10.1007/s11145-010-9238-y

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. doi:10.1016/S0022-5371(80)90312-6

Daneman, M., & Merikle, P.M. (1996). Working memory and comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. doi:10.3758/BF03214546

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks: Are they related with reading comprehension ability? *Learning and Individual Differences*, *12*(2), 131–143. doi:10.1016/S1041-6080(01)00033-4

De Beni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *Quarterly Journal of Experimental Psychology*, *51*(2), 305–320.
Duncan, J. (2010). *How intelligence happens*. New Haven, CT: Yale University Press.
Elosúa, M.R., García-Madruga, J.A., Gutiérrez, F., Luque, J.L., & Gárate, M. (1997). Un estudio sobre las diferencias evolutivas en la memoria operativa: ¿capacidad o eficiencia? [A study on developmental differences in working memory: Capacity or efficiency?]. *Estudios de Psicología*, *18*(2), 15–27.
Elosúa, M.R., García Madruga, J.A., Vila, J.O., Gómez-Veiga, I., & Gil, L. (in press). Improving reading comprehension: From metacognitive intervention on strategies to the intervention on working memory's executive processes. *Universitas Psychologica*.
Engle, R.W. (2001). What is working memory capacity? En H.L. Roediger, J.S. Naime, I. Neath, & A.M. Supremant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowde* (pp. 297–314). Washington, DC: American Psychological Association.
Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. doi:10.1111/1467-8721.00160
Engle, R.W., & Oransky, N. (1999). Multi-store versus dynamic models of temporary storage in memory. In R.J. Sternberg (Ed.), *The nature of cognition* (pp. 515–555). Cambridge, MA: MIT Press.
Engle, R.W., Tuholski, S.W., Laughlin, J., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable model approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331. doi:10.1037/0096-3445.128.3.309
Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245. doi:10.1037/0033-295X.102.2.211
Francis, D.J., Snow, C.E., August, D., Carlson, C.D., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *Scientific Studies of Reading*, *10*(3), 301–322. doi:10.1207/s1532799xssr1003_6
Friedman, N.P., Miyake, A., Corley, R.P., Young, S.E., DeFries, J.C., & Hewitt, J.K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179. doi:10.1111/j.1467-9280.2006.01681.x
García Madruga, J.A. (2006). *Lectura y conocimiento* [Reading and knowledge]. Barcelona, Spain: Paidos-UNED.
García-Madruga, J.A., Elosúa, M.R., Gutiérrez, F., Gárate, M., & Luque, J.L. (1999). *Comprensión lectora y memoria operative: Aspectos evolutivos e instruccionales* [Reading comprehension and working memory: Developmental and instructional issues]. Barcelona, Spain: Paidós.
García-Madruga, J.A., Gutiérrez, F., Carriedo, N., Luzón, J.M., & Vila, J.O. (2007). Checking the role of central executive in propositional reasoning. *Thinking & Reasoning*, *13*(4), 370–393.
García Madruga, J.A., Martín Cordero, J.I., Luque, J.L., & Santamaría, C. (1992). Teaching active text processing strategies. In B. van Hout-Wolters & W. Schnotz (Eds.), *Text comprehension from different perspectives* (pp. 183–200). Amsterdam: Swets-Zeitlinger.
Gaskins, I.W., Satlow, E., & Pressley, M. (2007). Executive control of reading comprehension in the elementary school. In L. Meltzer (Ed.), *Executive function in education: From theory to practice* (194–215). New York: Guilford.
Gathercole, S.E., & Baddeley, A.D. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.
Gathercole, S.E., Lamont, E., & Alloway, T.P. (2006). Working memory in the classroom. In S. Pickering (Ed.), *Working memory and education* (pp. 219–240). San Diego, CA: Academic. doi:10.1016/B978-012554465-8/50010-7
Goff, D.A., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing*, *18*(7–9), 583–616. doi:10.1007/s11145-004-7109-0
Gutiérrez-Martínez, F., García-Madruga, J.-A., Carriedo, N., Vila, J.-O., & Luzón, J.-M. (2005). Dos pruebas de amplitud de memoria operativa para el razonamiento [Two tests of working memory span for reasoning]. *Cognitiva*, *17*(2), 183–207. doi:10.1174/0214355054739255
Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component process of reading comprehension. *Journal of Educational Psychology*, *93*(1), 103–128. doi:10.1037/0022-0663.93.1.103
Hannon, B., & Daneman, M. (2004). Shallow semantic processing of text: An individual-differences account. *Discourse Processes*, *37*(3), 187–204. doi:10.1207/s15326950dp3703_1
Holmes, J., Gathercole, S.E., & Dunning, D.L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*(4), F9–F15. doi:10.1111/j.1467-7687.2009.00848.x
Jaeggi, S.M., Buschkuehl, M., Jonides, J., & Perrig, W.J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(19), 6829–6833. doi:10.1073/pnas.0801268105
Jaeggi, S.M., Studer-Luethi, B., Buschkuehl, M., Su, Y.-F., Jonides, J., & Perrig, W.J. (2010). The relationship between n-back performance and matrix reasoning—implications for training and transfer. *Intelligence*, *38*(6), 625–635. doi:10.1016/j.intell.2010.09.001
Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. doi:10.1037/0033-295X.99.1.122
Kane, M.J., & Engle, R.W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671. doi:10.3758/BF03196323
Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W., & Engle, R.W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. doi:10.1037/0096-3445.133.2.189
Kaufman, A.S., & Kaufman, N.L. (2000). *K-BIT, test breve de inteligencia de Kaufman* [K-BIT, the brief test of intelligence by Kaufman]. New York: PsychCorp.
Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, MA: Cambridge University Press.
Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*(7), 317–324. doi:10.1016/j.tics.2010.05.002
Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *24*(6), 781–791. doi:10.1076/jcen.24.6.781.8395
Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, *14*(4), 389–433. doi:10.1016/S0160-2896(05)80012-1
Melby-Lervåg, M., & Hulme, C. (2012). Is working memory training effective? A meta-analytic review. *Developmental Psychology*. Advance online publication. doi:10.1037/a0028228
Meltzer, L., Pollica, L., & Barzillai, M. (2007). Executive function in the classroom: Embedding strategy instruction into daily teaching practices. In L. Meltzer (Ed.), *Executive function in education: From theory to practice* (pp. 165–193). New York: Guilford.

Miller, G.A. (1977). Practical and lexical knowledge. In P.N. Johnson-Laird & P.C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 400–410). Cambridge, UK: Cambridge University Press.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. doi:10.1006/cogp.1999.0734

Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 442–481). New York: Cambridge University Press.

Morris, N., & Jones, D.M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*(2), 111–121. doi:10.1111/j.2044-8295.1990.tb02349.x

Oakhill, J.V., & Cain, K. (2007). Introduction to comprehension development. In K. Cain & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language* (pp. 3–40). New York: Guilford.

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H.-M. (2005). Working memory and intelligence—their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 61–65. doi:10.1037/0033-2909.131.1.61

Orjales, I., & García-Madruga, J.A. (2010). *Prueba de analogías para primaria* [Analogies span test for primary school]. Unpublished manuscript.

Orjales, I., García-Madruga, J.A., & Elosúa, M.R. (2010). *Prueba de amplitud lectora para primaria* [Reading span test for primary school]. Unpublished manuscript.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition*, *29*(2), 344–354. doi:10.3758/BF03194929

Rueda, M.R., Posner, M.I., & Rothbart, M.K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology*, *28*(2), 573–594. doi:10.1207/s15326942dn2802_2

Rueda, M.R., Rothbart, M.K., McCandliss, B.D., Saccomanno, L., & Posner, M.I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(41), 14931–14936. doi:10.1073/pnas.0506897102

Savage, R., Cornish, K., Manly, T., & Hollis, C. (2006). Cognitive processes in children's reading and attention: The role of working memory, divided attention, and response inhibition. *British Journal of Psychology*, *97*(3), 365–385. doi:10.1348/000712605X81370

Seigneuric, A., Ehrlich, M.-F., Oakhill, J., & Yuill, N. (2000). Working memory resources and children's reading comprehension. *Reading and Writing*, *13*(1/2), 81–103. doi:10.1023/A:1008088230941

Shipstead, Z., Redick, T.S., & Engle, R.W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628–654. doi:10.1037/a0027473

Snow, C., & Sweet, A.P. (2003). Reading for comprehension. In A.P. Sweet & C. Snow (Eds.), *Rethinking reading comprehension* (pp. 1–11). New York: Guilford.

Swanson, H.L., Howard, C., & Saez, L. (2006). Components of working memory that are related to poor reading comprehension and word recognition performance in less skilled readers. *Journal of Learning Disabilities*, *39*(3), 252–269.

Swanson, H.L., & Howell, M. (2001). Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. *Journal of Educational Psychology*, *93*(4), 720–734. doi:10.1037/0022-0663.93.4.720

Thorell, L.B., Lindqvist, S., Nutley, S.B., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*(1), 106–113. doi:10.1111/j.1467-7687.2008.00745.x

Unsworth, N., & Engle, R.W. (2005). Working memory capacity and fluid abilities: Examining the correlation between Operation Span and Raven. *Intelligence*, *33*(1), 67–81. doi:10.1016/j.intell.2004.08.003

Unsworth, N., & Engle, R.W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–132. doi:10.1037/0033-295X.114.1.104

van Oostendorp, H., & Goldman, S.R. (Eds.). (1999). *The construction of mental representations during reading*. Mahway, NJ: Erlbaum.

Vukovic, R.K., & Siegel, L.D. (2006). The role of working memory in specific reading comprehension difficulties. In T.P. Alloway & S.E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 89–112). New York: Psychology.

Wagoner, G.A. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, *18*(3), 328–346. doi:10.2307/747392

Whitney, P., Arnett, P.A., Driver, A., & Budd, D. (2001). Measuring central executive functioning: What's in a reading span? *Brain and Cognition*, *45*(1), 1–14. doi:10.1006/brcg.2000.1243

Submitted July 30, 2012
Final revision received December 21, 2012
Accepted December 31, 2012

JUAN A. GARCÍA MADRUGA is a professor in the Department of Developmental and Educational Psychology at the Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain; e-mail jmadruga@psi.uned.es.

MARÍA ROSA ELOSÚA is a professor in the Department of Basic Psychology at the UNED, Madrid, Spain; e-mail melosua@psi.uned.es.

LAURA GIL is a contractual professor in the Department of Developmental and Educational Psychology at the University of Valencia, Spain; e-mail laura.gil@uv.es.

ISABEL GÓMEZ-VEIGA is a contractual professor, **JOSÉ ÓSCAR VILA** an assistant professor, **ISABEL ORJALES** a contractual professor, and **ANTONIO CONTRERAS** an assistant professor in the Department of Developmental and Educational Psychology at the UNED, Madrid, Spain; e-mail igveiga@psi.uned.es, jovila@psi.uned.es, iorjales@psi.uned.es, and acontreras@psi.uned.es.

RAQUEL RODRÍGUEZ is an assistant professor in the Department of Methodology of Behavioral Sciences at the UNED, Madrid, Spain; e-mail rrodriguez@psi.uned.es.

MARÍA ÁNGELES MELERO is a professor in the Department of Developmental and Educational Psychology at the University of Cantabria, Santander, Spain; e-mail maria.melero@unican.es.

GONZALO DUQUE is a doctoral student in the Department of Developmental and Educational Psychology at the UNED, Madrid, Spain; e-mail gonzaloduque@bec.uned.es.