British Journal of Educational Psychology (2014), 84, 194–210

The British Psychological Society

www.wileyonlinelibrary.com

Improving reading comprehension in reading and listening settings: The effect of two training programmes focusing on metacognition and working memory

Barbara Carretti*, Nadia Caldarola, Chiara Tencati and Cesare Cornoldi

Department of General Psychology, University of Padova, Italy

Background. Metacognition and working memory (WM) have been found associated with success in reading comprehension, but no studies have examined their combined effect on the training of reading comprehension. Another open question concerns the role of listening comprehension: In particular, it is not clear whether training to improve reading comprehension must necessarily be based on processing written material or whether, as suggested in a recent study by Clarke *et al.* (2010, *Psychol. Sci., 21*, 1106), a programme based on verbal language could also be effective.

Aims. The study examined the feasibility of improving text comprehension in school children by comparing the efficacy of two training programmes, both involving metacognition and WM, but one based on listening comprehension, the other on reading comprehension.

Participants. The study involved a sample of 159 pupils attending eight classes in the fourth and fifth grades (age range 9–11 years).

Method. The listening and reading programmes focused on the same abilities/processes strictly related to text comprehension, and particularly metacognitive knowledge and control, WM (*per se* and in terms of integrating information in a text). The training programmes were implemented by school teachers as part of the class's normal school activities, under the supervision of experts. Their efficacy was compared with the results obtained in an active control group that completed standard text comprehension activities.

Results. Our results showed that both the training programmes focusing on specific text comprehension skills were effective in improving the children's achievement, but training in reading comprehension generated greater gains than the listening comprehension programme.

Conclusions. Our study suggests that activities focusing specifically on metacognition and WM could foster text comprehension, but the potential benefit is influenced by the training modality, that is, the Reading group obtained greater and longer-lasting improvements than the Active control or Listening groups.

*Correspondence should be addressed to Barbara Carretti, via Venezia 8, 35131 Padova, Italy (email: barbara.carretti@unipd.it).

DOI:10.1111/bjep.12022

Text comprehension is considered a complex cognitive ability that draws on a variety of language skills, including word-level lexical skills such as word-reading efficiency and vocabulary knowledge, and sentence-level skills such as knowledge of grammatical structure. Higher-order text-processing skills are also needed to understand a text, such as inference generation, comprehension monitoring, and working memory (WM) capacity (e.g., Cain, Oakhill, & Bryant, 2004; Oakhill & Cain, 2012). Higher-order skills are involved in text comprehension to enable the reader to make the integrative and inferential links needed to construct a meaning-based representation of the text, or what is called the *situational model* (Kintsch, 1998; van den Broek, 2010).

It has also been demonstrated that these higher-order factors help to explain reading comprehension performance, over and above the contribution of basic language skills, such as word reading. For example, Cain *et al.* (2004) demonstrated that WM capacity and specific component skills of comprehension (e.g., inference and integration skills, reading monitoring, and knowledge of story structure) explain a unique variance in reading comprehension between the ages of 8 and 11 years, after the contributions of word-reading skill and verbal ability have been taken into account (see also Oakhill & Cain, 2012).

The marginal role of reading decoding and associated abilities in reading comprehension has also emerged from the analysis of individual differences in comprehension. In a 4-year longitudinal study, Nation, Cocksey, Taylor, and Bishop (2010) showed that poor and good comprehenders (followed up from the age of 5 to 8 years) were comparable in terms of decoding and phonological skills, but the poor comprehenders had persistent oral language weaknesses involving expressive and receptive language, grammatical understanding, and listening comprehension (for similar results, see Catts, Adolf, & Ellis Weismer, 2006).

Summarizing the results emerging from normal development and individual differences analysis, it could be argued that reading comprehension is associated with good semantic language skills (e.g., vocabulary knowledge, morpho-syntactic skills) and good oral comprehension skills. Other cognitive (related with WM capacity and inference making) and metacognitive factors (knowledge and control) may contribute to good comprehension too (Hulme & Snowling, 2009). In fact, numerous studies examined the relevance of cognitive and metacognitive factors in reading comprehension also in the context of training studies. It is worth noting, however, that most of these studies considered each factor in isolation. The results are generally very encouraging as they show positive outcomes after training programmes focusing on metacognitive knowledge and control (e.g., Lucangeli, Galderisi, & Cornoldi, 1995) or on the ability to draw inferences and integrate different parts of a text (e.g., Yuill & Oakhill, 1991). There are also data suggesting that such positive effects are maintained, as in the report from Berkeley, Mastropieri, and Scruggs (2011), who analysed the efficacy of a strategy-instruction programme associated with an attributional retraining on reading outcomes for 7th, 8th, and 9th graders with learning and other mild disabilities.

Few studies have analysed the possibility of improving reading comprehension by means of WM training, and the results in children are contradictory (e.g., Dahlin, 2011; Swanson, Kehler, & Jerman, 2010), whereas more consistently positive results have been reported in older adults (e.g., Carretti, Borella, Zavagnin, & De Beni, 2012; Richmond, Morrison, Chein, & Olson, 2011). A recent meta-analysis by Melby-Lervåg and Hulme (2013) raised doubts about the efficacy of training programmes focusing on WM, in terms of their transfer and maintenance effects. It is also not clear whether WM training is applicable to educational settings, because most of the studies analysing the effects of WM

training in children involved individual not class sessions, which were not part of the normal school activities (e.g., Loosli, Buschkuehl, Perrig, & Jaeggi, 2012). In other words, although increasing attention has been paid to WM processes affecting learning achievement (e.g., Alloway, 2006), few studies have attempted to enhance WM in activities suitable for including in a school schedule. An exception is the study by Madruga *et al.* (2013), which embedded verbal WM training activities in a group intervention on reading comprehension, in which WM was trained using tasks that involved following complex instructions, variants of the Daneman and Carpenter Listening Span test (1980), and texts requiring the updating of information: Using this approach, the authors succeeded in ameliorating both WM and reading comprehension performance.

Objectives of the study

In the light of findings in the literature, the first main aim of our study was to analyse the efficacy of two training programmes focusing on specific skills/abilities associated with text comprehension (i.e., metacognition, verbal WM and the ability to integrate information) that have emerged as fundamental to our understanding of reading comprehension processes in typically developing children (e.g., Cain *et al.*, 2004) and those with specific difficulties in this area (e.g., Hulme & Snowling, 2009).

In addition to metacognition and WM, we focused on the ability to integrate different information because of its close connection with WM, which helps readers to keep different pieces of information from the text actively in mind and integrate them in a coherent mental representation (Spooner, Gathercole, & Baddeley, 2006), our aim being to facilitate transfer effects of the WM activities on reading comprehension activities.

A second aim of our study was to compare the efficacy of two training delivery modalities, oral versus written. Based on the close link demonstrated between oral and written comprehension at various ages (e.g., in children [Berninger & Abbott, 2010] and in young adults [Gernsbacher, Varner, & Faust, 1990]), and assuming that oral comprehension and written comprehension rely on much the same linguistic skills, training based on a verbal modality (listening) might be expected to improve performance in reading comprehension too. Aarnoutse, Van Den Bos, and Brand-Gruwel (1998) analysed this issue by delivering training orally to 9- to 11-year-old children classified according to their reading decoding and listening comprehension levels, and assessing the effect of the training on reading and listening comprehension measures. Participants attended 20 sessions conducted by trained experimenters in small groups (of six participants each), while a control group attended 20 regular reading comprehension lessons, without receiving any instruction on comprehension strategies, and the sessions were carried out with the whole class. The results showed a better use of strategies by the experimental group (not by the control group) when the children heard or read a text, but this did not produce any changes in the children's performance in either listening or reading comprehension tasks (see also Aarnoutse, Brand-Gruwel, & Oduber, 1997; Brand-Gruwel, Aarnoutse, & Van Den Bos, 1998; van den Bos, Aarnoutse, & Brand-Gruwel, 1998).

On the other hand, Clarke, Snowling, Truelove, and Hulme (2010) found a positive influence of oral language training on reading comprehension in a group of poor comprehenders (i.e., children with adequate decoding skills, but difficulties in text comprehension). Their study compared the efficacy of three text comprehension training programmes using more classic reading activities (reading comprehension programme), or activities designed to improve oral language competence (oral language programme),

or a combination of reading and oral language activities. The authors demonstrated that children with reading comprehension difficulties benefited more from the oral language programme than from the other two, and the improvement recorded immediately after the training persisted at a follow-up 11 months later, offering support for the hypothesis that reading comprehension problems are particularly related to difficulties in oral language (see Hulme & Snowling, 2011). The written and oral training programmes did not work on the same abilities/processes, however. In fact, the text reading comprehension programme mainly provided training on metacognition and inference, whereas the oral language programme focused on vocabulary and figurative language. It would therefore be interesting to see whether the efficacy of the oral training depended on specific characteristics of the children involved or on the abilities that were trained, and whether the results could be replicated in other settings. A strength of the Clarke *et al.* study lies in that they trained children with reading comprehension difficulties using activities focusing on oral language skills, which are likely to be more motivating for children who are struggling with written texts. It is therefore important to establish whether the same would apply to normal populations in the context of normal school activities. In the present study, training was administered to normally developing children in either an oral or a written format, focusing on certain cognitive processes, such as metacognition (e.g., Baker & Brown, 1984) and WM capacity (e.g., Daneman & Merikle, 1996), which have been found to correlate with reading comprehension and might be particularly appropriate for an oral training scheme too because they do not demand the use of written material.

To clarify the effect of the training delivery modality, the activities involved either reading or listening, but they were substantially the same to avoid differences in the results being inflated by other factors. The two training programmes were compared with a third programme based on materials already used in the school to promote reading comprehension, but prepared so as to make the programme seem new and innovative to the teachers; these activities typically involved reading one or more texts and then answering questions, producing summaries, and so on.

The 22 training sessions involved in our study were conducted by the children's teachers, who were not informed about the study design and who had no prior experience of metacognitive and WM activities. To contain the problem relating to the teachers' limited familiarity with the specific procedures involved in our project, we provided the teachers with 4 hr of training on how to implement the activities in the classroom, and we met them at various times while the programme was underway to ensure that the activities were implemented properly.

Method

Participants

The project involved an initial sample of 185 students attending eight classes in the fourth and fifth grades (age range 9–11 years) at six schools in and around Padova (north-eastern Italy). The eight classes were similar in terms of the children's gender and abilities (see below), teacher-reported sociocultural level, and their teachers' experience, motivation, and expertise. The sociocultural level of the classes reflected the characteristics of the population in the areas in and around Padova: 66% of the children came from families with a middle-to-high sociocultural level, and 34% from families with a low sociocultural level. The teachers were all women aged between 40 and 55, and they all had at least 15 years of teaching experience. Three of the eight classes completed our reading comprehension

Table 1. Characteristics of the three groups

		Vocabulary		Lexical decision		Mental rotation	
	N	*M*	*SD*	*M*	*SD*	*M*	*SD*
Reading group	45	22.59	6.49	31.47	8.91	14.36	4.08
Listening group	28	21.04	7.72	35.67	11.49	14.97	4.25
Active control group	57	23.07	6.06	32.37	8.38	13.93	3.29

training (Reading group, 48% females), two were administered the listening comprehension training (Listening group, 38% females), and three had the standard reading comprehension lessons on the school curriculum (Active control group, 44% females).

Participants were initially assessed using a set of group-administered tasks measuring general and specific reading-related abilities. In particular, tasks were administered to assess their vocabulary knowledge and their ability to distinguish words from non-words (lexical decision) and manipulate spatial information. This was done to obtain some baseline information about the groups and ensure that they were comparable. The lexical decision test (Caldarola, Perini, & Cornoldi, 2012) was a paper-and-pencil task consisting of three lists of words and non-words (for a total of 120 stimuli): Participants were asked to identify as many non-words as they could in 2 min. The other tasks used were drawn from the Italian version of the primary mental abilities battery of vocabulary and spatial tests (Thurstone & Thurstone, 1963). The three groups did not differ in the above-mentioned measures (see Table 1).

Children with intellectual disabilities and those for whom Italian was their second language (18% of the total) took part in the activities but were excluded from the analysis.

Material

Pre-test/post-test assessment

Specific effects. *Metacognition* (adapted from De Beni & Pazzaglia, 1991): A questionnaire composed of nine items was used to assess metacognitive knowledge and control. The questionnaire assesses different aspects of metacomprehension, particularly text sensitivity, strategy use, and self-monitoring. Text sensitivity is considered in terms of the ability to distinguish between important and unimportant ideas, and awareness of the varying levels of text difficulty and the variety of text genre. Strategy use is assessed with questions measuring the efficacy of reading strategies and the appropriateness of their use according to the literary genre. The self-monitoring items involve participants identifying inconsistencies in a series of sentences. The score used was the total number of correct answers (maximum score = 13), awarding one point for correct answers. It must be noticed that, for four of the nine items, the score ranged from 0 to 2 depending on the level of correctness of the response. For example, in the item where respondents had to identify inconsistences in a series of 10 sentences, they were awarded two points if they answered correctly for at least five sentences, one point for 3–4 correct answers, no points for 1–2 correct answers.

Working memory. As updating information in WM seems particularly crucial to reading comprehension (e.g., Carretti, Cornoldi, De Beni, & Romanò, 2005), we decided to use a WM updating task adapted from the updating-following-a-relevant-criterion task proposed by Palladino, Cornoldi, De Beni, and Pazzaglia (2001). Four lists of nouns were presented to the child, who had to remember three items in each list according to the

criterion 'remember the smallest object(s) in each list', and in the right order of presentation. All the words were highly familiar to the children and referred to objects that were easy to compare for in terms of size. The number of correctly recalled words was the dependent variable (maximum score 18).

Integration skills. In this task, adapted from De Beni, Cornoldi, Carretti, and Meneghetti (2003), participants were asked to co-referentially correlate elements found in different parts of the text and/or illustrations, to connect synonyms and words relating to the same characters, and to connect information so as to attribute the correct meaning to words. After reading a text, participants answered 14 open-ended or multiple-choice questions. The number of correct answers was the dependent variable (maximum score 14).

Transfer effects on comprehension

Reading comprehension. The reading comprehension task was drawn from the Italian standardized battery for the assessment of reading decoding and reading comprehension in primary school (Cornoldi & Colpo, 2011). Following the instructions in the manual, different texts suited to the different school years and corresponding to the different time points in the study were used to assess reading comprehension at the pre-test, post-test, and follow-up sessions. This enabled us to use the corresponding normative data as a control measure of how reading comprehension skills are acquired in typically developing children without specific training. The children were asked to read the text silently and then answer questions; the passage remained available for them to reread or consult while they were answering. This procedure is adopted to limit the influence of memory on performance. The final score was calculated as the sum of the correct responses (maximum score 14).

Listening comprehension task. The task came from an Italian standardized battery for assessing listening comprehension (Carretti, Caldarola, Tencati, & Cornoldi, 2013) and consisted of a text read aloud by the experimenter, followed by 12 multiple-choice questions. To contain the influence of memory, the text was read in two parts and participants were asked the questions about the first part before moving on to the second. The final score was obtained from the sum of the correct responses (maximum score 12).

The teachers' training and supervision. Before starting to train the pupils, we met the teachers interested in using the materials for an introductory session lasting about 4 hr. The first part of the session was plenary. The programmes were presented without providing details about the experimental aims of the project, but particularly emphasizing the frequency of the sessions each week, the schedule of each session, and how to instruct the children to complete the activities. The teachers were also told that all the programmes were expected to benefit the children's reading comprehension skills. This first part took about 1 hr. Then, the teachers were quasi-randomly assigned to the three training programmes, and, for the remaining 3 hr, they were divided into three groups, and the specific activities for one of the training programmes were presented. Before starting to train the children, all the teachers were contacted again to clarify any doubts or problems and to hand over and jointly analyse the training materials.

After the first half of the training sessions had been completed, we met the teachers again to discuss any problems and ask for their opinions on the training activities. This was

carried out for all the teachers involved in the study, separately for each of the three training programmes. During the children's training, the teachers were monitored individually by two expert researchers, who also advised them as to the best way to conduct the sessions. This enabled us to assess the quality of delivery of the training.

Training activities. The three training programmes consisted of 22 sessions, conducted by the previously instructed teachers. Each session lasted about 1 hr and followed a fixed schedule, as shown in Figure 1, and there were two sessions a week. The session was held collectively, but children worked on their own. At the beginning of each session, pupils were given materials specifically designed for the activities, and after the teacher had introduced the activities, they completed the various exercises, depending on the training programme concerned. To reinforce the effect of the training, two of the sessions (sessions 8 and 15) were used to revise the activities of previous sessions; these sessions were also scheduled for the Active control group.

Table 2 summarizes the topics of each session for the two experimental groups (reading and listening conditions). To train metacognitive skills, participants were asked in the first five sessions (1–5) to reflect on the goals of reading and listening using different materials and examples. Then, three reading or listening strategies were introduced in the next three sessions (6, 7, and 9), emphasizing that the choice of a given strategy depends on the reading/listening goals. Sessions 10–14 involved activities devoted to improving the children's ability to monitor their level of comprehension. Then, in the last part of the training (sessions 16–22), participants were taught to recognize the text genre and to reflect on its implications for their use of comprehension strategies and on the importance of the title for predicting the content of a text.

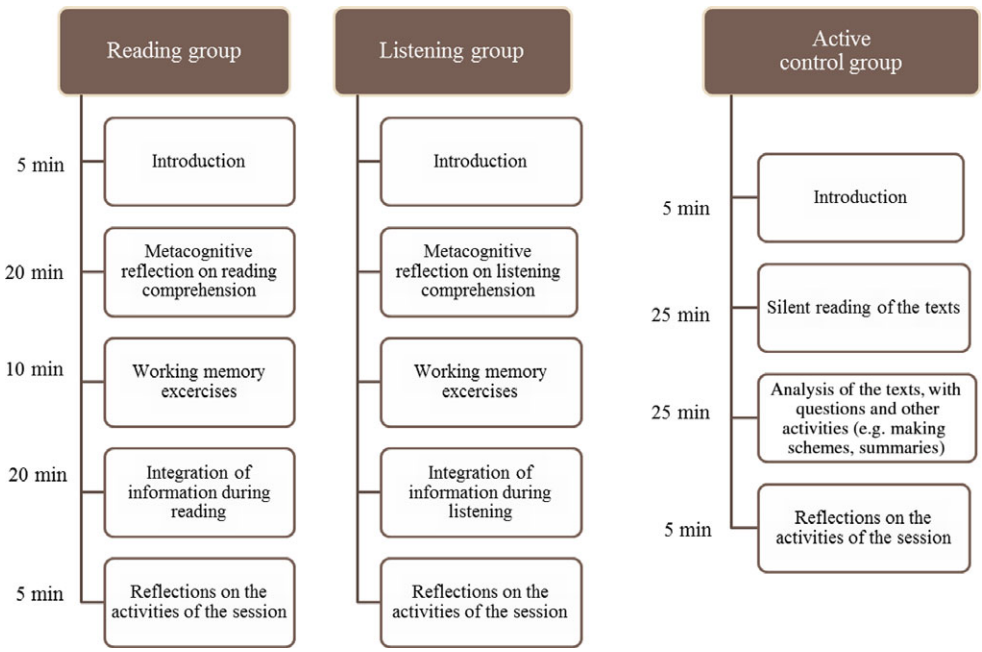

Figure 1. Schedule of each session by group.

Table 2. Detailed aims for the reading and listening training programme depending on the component (metacognition, working memory, and integration skills)

	Metacognition	Working memory	Integration skills
Session 1	Reading/Listening goals	LST without secondary task	Integration between different part of the text
Session 2	Reading/Listening goals	LST without secondary task	Integration between different part of the text
Session 3	Reading/Listening goals	LST without secondary task	Integration between text and pictures
Session 4	Reading/Listening goals	LST without secondary task	Integration between text and pictures
Session 5	Reading/Listening goals	LST without secondary task	Integration between text and pictures
Session 6	Reading/Listening strategies	LST with secondary task	Integration between text and pictures
Session 7	Reading/Listening strategies	LST with secondary task	Integration between two texts
Session 8	Revision of previous sessions	LST with secondary task	Revision of previous sessions
Session 9	Reading/Listening strategies	Recall of words with secondary task	Integration on the basis of the meaning
Session 10	Monitoring	Recall of words with secondary task	Integration on the basis of the meaning
Session 11	Monitoring	Recall of words with secondary task	Integration on the basis of the meaning
Session 12	Monitoring	Recall of words with secondary task	Integration on the basis of the meaning
Session 13	Monitoring	Recall of words with secondary task	Integration on the basis of the meaning
Session 14	Monitoring	Recall of words with secondary task	Integration on the basis of the meaning
Session 15	Revision of previous sessions	Updating with words	Revision of previous sessions
Session 16	Text sensitivity	Updating with words	Integration on the basis of the meaning
Session 17	Text sensitivity	Updating with words	Detecting relevant information and updating
Session 18	Text sensitivity	Updating with digits	Detecting relevant information
Session 19	Text sensitivity	Updating with words	Detecting relevant information
Session 20	Text sensitivity	Updating with words	Detecting relevant information
Session 21	Text sensitivity	Updating with digits	Detecting relevant information
Session 22	Text sensitivity	Updating with words	Detecting relevant information

LST = listening span test.

For the purposes of training WM, the activities were always presented in a listening format, that is, the teacher read the material (lists of sentences, words or digits) aloud and the pupils were asked to write the information they had to recall in a dedicated booklet. In the first sessions (1–4), adapting the procedure typical of the Listening span test proposed by Daneman and Carpenter (1980), participants were presented with increasing numbers of sentences (from 2 to 5) and asked to recall the last word in each sentence in serial order. In sessions 6–8, new sets of sentences were provided, but participants had to complete a secondary task, that is, to decide whether each sentence was true or false, recording this in the dedicated booklet. Adapting the task originally proposed by De Beni, Palladino, Pazzaglia, and Cornoldi (1998), sessions 9–15 involved participants being presented with an increasing number of word lists (2–5) and having to recall the last word in each list in serial order, while also putting a cross in the dedicated booklet whenever they heard an animal noun. In the last sessions (16–22), there were exercises that involved updating memorized information, using the same procedure as in the task administered at the pre-test and post-test points (i.e., having to recall the three smallest items in a list in serial order); in two of the sessions (18 and 21), however, the list contained digits instead of words.

For the part of the training that focused on the ability to integrate different information, the first sessions (from 2 to 7) trained the pupils to connect information within a given text, between a text and a picture, and between two different texts. In sessions 9–14, and 16, participants were asked to integrate information in order to construct a coherent mental model of the text, using narrative and expository texts. The last sessions (18–22) had the same objectives, but participants were asked to update information while reading or to identify the importance of the information in the text according to different aims.

Results

First, to identify any differences between the groups receiving the different types of training, separate ANOVAs were run on the groups' pre-test performance in all tasks, with group (Reading, Listening, and Active control) as the between-subjects factor. The results indicated that there were no baseline differences between the groups, with the exception of the number of items recalled in the updating task, $F(2, 123) = 3.42$, $\eta_p^2 = .05$, $p < .05$, in which the Active control group outperformed the Reading training group ($p < .001$). No other differences were significant.

To assess the effects of the training, the measures of interest for each task were analysed using three group (Reading, Listening, and Active control) × 2 Session (pre-test, post-test) mixed-design ANOVA, with group as the between-subjects factor and sessions as the repeated measures. Interactions were broken down using *post-hoc* pairwise comparisons with Bonferroni's correction at $p < .05$, adjusted for multiple comparisons. Descriptive statistics are given in Table 3 (for the specific effects) and Figure 1 (for the transfer effects).

Specific effects

Metacognition

The main effect of group was significant, $F(2, 130) = 4.59$ $\eta_p^2 = .07$ $p < .05$, with the Reading group performing better than the Active control group ($p < .05$), and so was the main effect of session, $F(1, 130) = 68.75$, $\eta_p^2 = .35$, $p < .001$, because post-test performance was better than at the pre-test stage (see Table 3). The Group × Session

Table 3. Means and standard deviations in the Metacognition, Updating, and Integration tests, observed before and after training in the three groups

	Metacognition		Working memory updating		Integration skills	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Reading group						
M	8.61	11.64	9.44	11.07	8.88	9.88
SD	3.34	3.36	2.05	2.12	2.93	2.81
Listening group						
M	8.14	11.02	9.60	9.96	8.08	9.72
SD	4.10	2.79	2.00	2.65	3.14	2.81
Active control group						
M	7.90	9.11	10.49	10.33	7.55	8.82
SD	2.73	3.01	2.82	2.84	2.49	2.69

interaction was significant too, $F(2, 130) = 5.05, \eta_p^2 = .07, p < .01$. Post-hoc comparisons showed that all the groups' performance improved from pre- to post-test (Reading group $p < .001$; Listening group $p < .001$; Active control group $p < .01$), but, at the post-test stage, the Reading and Listening groups also performed better than the Active control group ($p < .001$ and $p < .001$, respectively).

WM updating

The main effect of session was significant, $F(1, 120) = 4.23, \eta_p^2 = .03, p < .05$, given that post-test performance was better than at the pre-test point. The Group × Session interaction was significant, $F(2, 120) = 4.11, \eta_p^2 = .06, p < .05$. As shown in Table 3, the groups differed slightly at both pre- and post-test assessments, and only the two experimental groups improved. Post-hoc comparisons showed that this improvement was only statistically significant in the case of the Reading group ($p < .001$).

Integration skills

The main effect of session was significant, $F(1, 118) = 34.59, \eta_p^2 = .23, p < .001$, because post-test was better than pre-test performance for all groups (see Table 3). No other effects were significant.

Transfer effects on reading and listening comprehension

Figure 2 shows the standardized scores (calculated on the basis of the test norms) obtained by the three groups of children at pre-test, post-test, and follow-up. We first analysed the changes between pre-test and post-test. For both reading and listening comprehension, the groups differed very little at pre-test in that the groups had a near-perfectly average performance (the control group performed slightly below the normative mean), and all the groups' performance improved, as confirmed by the significant main effect of session in both reading comprehension, $F(1, 118) = 8.37$, $\eta_p^2 = .07$, $p < .01$, and listening comprehension, $F(1, 122) = 17.25$, $\eta_p^2 = .12$, $p < .001$, but the improvements differed in the three groups as demonstrated when we analysed the dimension of the changes.

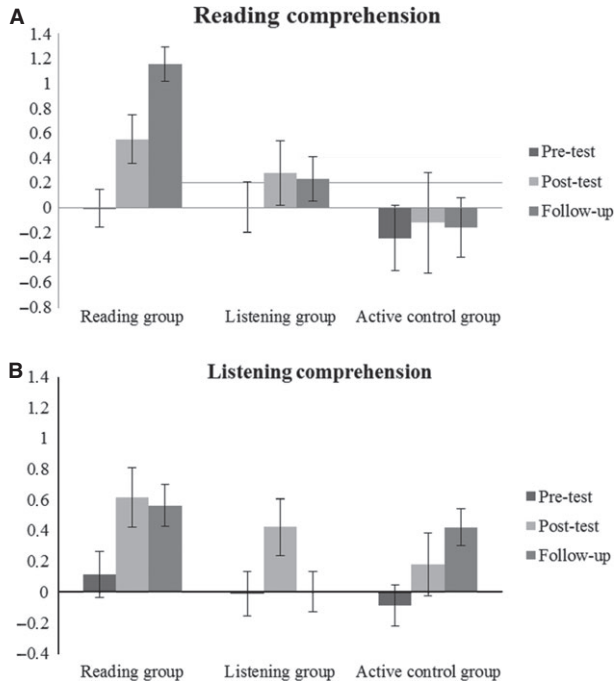

Figure 2. Transfer effects to reading comprehension (panel A, above) and listening comprehension (panel B, below) by group (bars represent standard errors).

To gain a better understanding of the range of training gains and transfer effects between pre- and post-test in the two experimental groups, we also calculated Cohen's d (1988), which expresses the effect size of the comparisons. When the gains from pre- to post-test were compared within each group in all the tests, the effect size was low in the Active control group. In contrast, it ranged from medium to large for the Reading group in all the measures considered; in the case of the Listening group, the effect size was null in the WM measure, while it ranged from small to large in the other measures (see Figure 3).

Analysing the improvement from pre-test to follow-up:[1] Eight months after completing the training, the subgroups of children who could be reached again were administered follow-up tests, and their performance was compared with the three groups' pre-test performance. The Reading group's follow-up performance was higher than at the pre-test stage for both the Reading comprehension task, $t(43) = 3.85, p < .001$, and the Listening comprehension task, $t(43) = 3.14, p < .01$. In the Listening group, this was only true for the Reading comprehension task, $t(17) = 2.16\ p < .05$; and the Control group performed better at follow-up in the Listening comprehension task, $t(55) = 2.93, p < .01$.

[1] *The dropout rate differed between the three groups: 14 children in the Reading group could not be reached, 10 in the Listening group, and two in the Active control group. There were several reasons for the high dropout rate in the first two groups: Some of the children changed schools and the teacher changed in one class, making it impossible to test the children. When the dropouts were compared with the other children, there were no differences in their pre-test performance in the Reading and Active control groups, but in the Listening group, there were several differences (in both directions): Dropouts performed worse in the Vocabulary task* $F(1, 26) = 8.06$, $p < .01$, *and better in the Mental rotation* $F(1, 27) = 7.04$, $p < .05$ *and reading comprehension tasks* $F(1, 25) = 6.74$, $p < .05$ *than the children who were also tested at follow-up. No differences emerged for the Lexical decision and Listening comprehension tasks. These results mean that the measures obtained at the follow-up in the Listening group need to be considered with caution.*

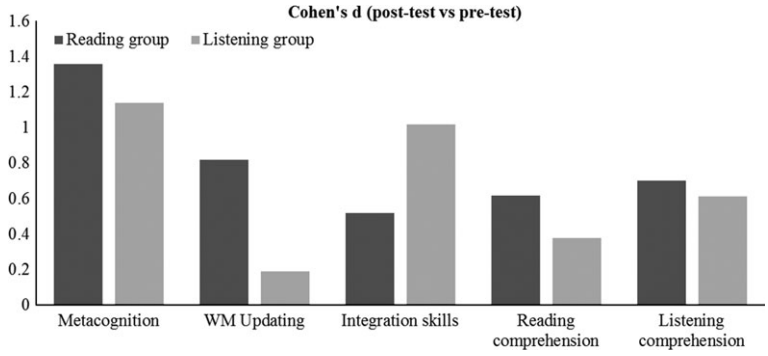

Figure 3. Effect size (d) for tasks measuring specific and transfer effects of the training by group.

Relationship between training-induced gains in specific measures and performance in reading and listening comprehension tasks

Although the number of participants within each training group was limited, we analysed the relationship between the training-induced gains in specific tasks and performance in reading and listening comprehension tasks in the two experimental training groups (i.e., in the reading and listening modalities). Gains were computed as the difference between the post-test and pre-test performance divided by the standard deviation at the pre-test for the whole group (see Loosli *et al.*, 2012). The correlation between the above-mentioned measures clearly showed a significant correlation between gains in WM and performance in the listening comprehension task at the post-test stage ($r = .461$, $p < .001$) for the Reading group. No further correlations were significant.

Discussion and conclusion

The goals of this study were to assess the efficacy of two training programmes focusing on metacognition, WM, and integration skills, using two different approaches: One in which training was provided by means of reading activities and the other using the same material administered orally. The two training programmes were compared with a third involving standard school activities for teaching reading comprehension.

Several studies have demonstrated that not only general skills but also difficulties in reading comprehension relate to inadequate metacognitive knowledge and control (e.g., Cataldo & Cornoldi, 1998; Cataldo & Oakhill, 2000), as well as poor WM capacity (e.g., Carretti, Borella, Cornoldi, & De Beni, 2009). For this reason, and to facilitate transfer effects on general comprehension skills, our training programmes combined activities focusing on WM with those for developing integration skills. Another important feature of the present study was that teachers were directly involved in the training activities, and they conducted the sessions during normal classroom activities. Our results should therefore hold for the typical school setting.

As concerns the specific effect of the training, the activities presented using both a reading and a listening format beneficially influenced metacognition, with a large effect size. In fact, both the Reading and the Listening groups improved in their performance from pre-test to post-test, and, more importantly, they outperformed the Active control group at the post-test stage. The Active control group also improved to some degree, confirming that traditional activities for improving children's reading comprehension may produce a benefit too.

An improvement in WM was only seen in the Reading group, and its effect size was medium; the Listening group showed no such benefit. The improvement in WM obtained by the Reading group, measured in terms of standardized gains (see Loosli *et al.*, 2012), correlated with the group's performance in the Listening comprehension task at the post-test stage, offering some preliminary evidence of the causal role of an improved WM. Thus, although the present findings support the feasibility of enhancing WM, as other studies on children have already demonstrated (e.g., Holmes, Gathercole, & Dunning, 2009; Loosli *et al.*, 2012), the findings that only one of the two groups receiving the WM training (based on a listening modality for both groups) improved and that only one correlation was significant suggest that caution is warranted. One reason why only the Reading group's WM improved could be that the format of the WM exercises generated more interest and attention in the Reading group because the setting of the training changed, whereas the whole session was conducted in the same modality for the Listening group, so the WM exercises were not seen as differing from the other activities. It may also be that the Listening group's auditory WM had already been overloaded by the previous activities on metacognition, so these children were possibly more tired and less attentive. Further studies might include qualitative or quantitative evaluations on this issue.

In the case of integration skills, all three groups' performance improved, including the Active control group, indicating that the standard school activities focusing on reading comprehension also promote these skills.

The transfer effects on comprehension seem particularly interesting. The most important effects were seen in the Reading group, with improvements in both reading comprehension and listening comprehension at the post-test assessments (as shown by the effect size indexes). This means that the comprehension training provided using a reading modality produced positive effects not only in the same modality (reading), but also in a different but correlated modality (listening). In the case of the Listening group, we observed a transfer to listening comprehension at the post-test point, but no effect on reading comprehension. It should be noted that the z-scores for the reading and listening comprehension tasks were computed using national norms so the groups' improvement could be compared with the expected performance of large groups of children observed at the same point in the school year, who served as a sort of passive control. This comparison showed that the trained groups' performance (that of the Reading group in particular, less so for the Listening group) was always above the expected average performance.

The data we collected at the follow-up stage should be considered with caution because they only concern a part of our sample of children, and it was impossible to control for other classroom activities carried out by the teachers in the intervening period. It is nonetheless worth noting that, after 8 months, the Reading group substantially maintained the gains seen at the post-test assessments, while the Listening group's performance dropped back to pre-test levels. It may be that this was due to a combined improvement in metacognition and WM experienced by the Reading group; this would support the conviction that these aspects are crucial to comprehension (see Cain *et al.*, 2004), as also suggested by our results as a whole, irrespective of the training modality used. Future studies should systematically examine whether improvements in the underlying processes (metacognition, WM and information integration) coincide with improvements in reading and listening comprehension. A preliminary finding to suggest that this is so lies in the fact that the improvement in the Reading group's WM correlated positively with performance in the listening comprehension task at the post-test stage.

This result was isolated, however, and should be considered with caution, also in view of the relatively small numbers of children in each group.

To sum up, our results show that specific programmes may generally produce advantages in comprehension and the underlying cognitive processes. The differences seen between the experimental groups and the Active control group seem important, however. In particular, we found larger gains for the reading comprehension programme than for the other two programmes. Our listening comprehension training produced some benefits, but they were restricted to the same modality and were not maintained at the follow-up (although our follow-up findings could be biased by the particular characteristics of the children who dropped out). The impression generated by our results, that is, that an approach based on reading activities is more effective in promoting reading and listening comprehension than one based on a listening modality (in our sample and under our conditions at least) partially contrasts with the report from Clarke *et al.* (2010). Several important differences could explain the different outcomes between the Clarke *et al.* study and our own, such as the children's ages (8–9 vs. 9–11), the setting (small groups of children vs. classes), and the trainers' and children's personal characteristics. In particular, our sample included all the children in the classes, whereas the Clarke study only involved children with difficulties, who were presumably less motivated to use written material at school. We also found that the teachers involved were more at ease with activities on written texts than on listening comprehension exercises. Another aspect to consider concerns the language because differences between English and Italian could explain the advantage of the reading programme over the listening-based training in our sample. Italian fourth and fifth graders already have fully automated decoding skills, unlike children learning to read in a language with a deep orthography (like English; e.g., Aro & Wimmer, 2003), so the Italian children do not encounter particular difficulties in the decoding process.

This last aspect is also important when we consider the possible role of reading decoding in explaining the strength of the transfer effects. It may be that the advantage of the Reading programme stemmed from the fact that it trained the children's reading decoding skills as well as their reading comprehension skills. In other words, the positive outcomes in the Reading group may be amplified by an improvement in their decoding skills. This might partially account for the more limited effects in the Listening group. Italian fourth- and fifth-grade students are usually fluent readers, however, and it has been demonstrated that the contribution of reading decoding in explaining reading comprehension is modest in children learning to read in languages with transparent orthography at the age considered here (e.g., Florit & Cain, 2011).

In conclusion, despite some limitations (such as the small number of participants, which prevented us from analysing any effects of the teachers' characteristics or the contribution of the separate components in producing the gains observed), our study contributes to our understanding of the processes involved in reading comprehension. We demonstrated the effectiveness of activities with a combined effect on metacognition and WM, confirming that these two factors are closely related to reading comprehension adequacy. Our findings also indicate that the potential benefit of text comprehension training depends on the modality used, because the Reading group obtained larger and longer-lasting improvements than the Active control or Listening groups.

These results seem useful also for the purposes of day-to-day school activities, that is, from an educational standpoint, it seems noteworthy that training provided by teachers and focusing on the processes underlying reading comprehension could have substantial

positive effects, meaning that specific activities could be included routinely in the school curricula.

Acknowledgements

The present study was conducted with the support of a grant from the University of Padova awarded to Cesare Cornoldi (Progetto Strategico No STPD08HANE_002).

References

Aarnoutse, C., Brand-Gruwel, S., & Oduber, R. (1997). Improving reading comprehension strategies through listening. *Educational Studies*, *23*, 209–227. doi:10.1080/0305569970230205

Aarnoutse, C., Van Den Bos, K. P., & Brand-Gruwel, S. (1998). Effects of listening comprehension training on listening and reading. *The Journal of Special Education*, *32*, 115–126. doi:10.1177/002246699803200206

Alloway, T. P. (2006). How does working memory work in the classroom? *Educational Research and Reviews*, *1*, 134–139. Retrieved from http://www.academicjournals.org/ERR/PDF/Pdf2006/Jul/Alloway.pdf

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*, *24*, 621–635. doi:10.1017/S0142716403000316

Baker, L., & Brown, A. L. (1984). Metacognition skills of reading. In D. Pearson, R. Barr, M. Kamil & P. Monsenthal (Eds.), *Handbook of reading research* (pp. 353–394). New York, NY: Longman.

Berkeley, S., Mastropieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities*, *44*, 18–32. doi:10.1177/0022219410371677

Berninger, V., & Abbott, D. (2010). Listening comprehension, oral expression, reading comprehension and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology*, *102*, 635–651. doi:10.1037/a0019319

Brand-Gruwel, S., Aarnoutse, C. A. J., & Van Den Bos, K. P. (1998). Improving text comprehension strategies in reading and listening settings. *Learning and Instruction*, *8*, 63–81. doi:10.1016/S0959-4752(97)00002-9

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*, 31–42. doi:10.1037/0022-0663.96.1.31

Caldarola, N., Perini, N., & Cornoldi, C. (2012). DLC: una prova di decisione lessicale per la valutazione collettiva delle abilità di lettura [DLC: A lexical decision task for the evaluation of reading decoding]. *Dislessia*, *9*, 89–104. Retrieved from http://rivistedigitali.erickson.it/dislessia/archivio/vol-9-n-1-2/

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

Carretti, B., Borella, E., Zavagnin, M., & De Beni, R. (2012). Gains in language comprehension relating to working memory training in healthy elderly adults. *International Journal of Geriatric Psychiatry.* Advance online publication, doi:10.1002/gps.3859

Carretti, B., Caldarola, N., Tencati, C., & Cornoldi, C. (2013). *Comprensione Orale – Test e trattamento [Oral comprehension – Test and treatment]*. Trento, Italy: Erickson.

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*, 45–66. doi:10.1016/j.jecp.2005.01.005

Cataldo, M. G., & Cornoldi, C. (1998). Self-monitoring in poor and good reading comprehenders and their use of strategy. *The British Journal of Developmental Psychology*, *16*, 155–165. doi:10.1111/j.2044-835X.1998.tb00915.x

Cataldo, M. G., & Oakhill, J. (2000). Why are poor comprehenders inefficient searchers? An investigation into the effects of text representation and spatial memory on the ability to locate information in text. *Journal of Educational Psychology*, *92*, 791–799. doi:10.1037/0022-0663,92.4.791

Catts, H. W., Adolf, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*, 278–293. doi:10.1044/1092-4388(2006/023)

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science*, *21*, 1106–1116. doi:10.1177/0956797610375449

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cornoldi, C., & Colpo, G. (2011). *Prove di lettura MT-2 per la scuola primaria [Test for reading assessment in the primary school]*. Firenze, Italy: Giunti OS.

Dahlin, K. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing*, *24*, 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*, 450–466. doi:10.1016/S0022-5371(80)90312-6

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

De Beni, R., Cornoldi, C., Carretti, B., & Meneghetti, C. (2003). *Nuova Guida alla Comprensione del Testo: Volume 1 [New guide for reading comprehension: Vol. 1]*. Trento, Italy: Erickson.

De Beni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *51A*, 305–320. doi:10.1080/713755761

De Beni, R., & Pazzaglia, F. (1991). *Lettura e metacognizione [Reading and metacognition]*. Trento, Italy: Erickson.

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*, 553–576. doi:10.1007/s10648-011-9175-6

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 430–445. doi:10.1037/0278-7393.16.3.430

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory children. *Developmental Science*, *12*, 9–15. doi:10.1111/j.1467-7687.2009.00848.x

Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. Chichester, UK: Wiley-Blackwell.

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science*, *20*, 139–142. doi:10.1177/0963721411408673

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *18*, 62–78. doi:10.1080/09297049.2011.575772

Lucangeli, D., Galderisi, D., & Cornoldi, C. (1995). Specific and general transfer effects following metamemory training. *Learning Disabilities Research and Practice*, *10*, 11–21.

Retrieved from http://ovidsp.tx.ovid.com/sp-3.8.1a/ovidweb.cgi?&S=LELGFPENNPDDAFKDNCOKOGFBFPEOAA00&Complete+Reference=S.sh.17%7c10%7c1

Madruga, J. A., Elosúa, M. R., Gil, L., Gómez-Veiga, I., Vila, J. O., Orjales, I., … Duque, G. (2013). Reading comprehension and working memory's executive processes: An intervention study in primary school students. *Reading Research Quarterly*, *48*, 155–174. doi:10.1002/rrq.44

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*, 270–291. doi:10.1037/a0028228

Nation, K., Cocksey, J., Taylor, J. S. H., & Bishop, D. V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry*, *51*, 1031–1039. doi:10.1111/j.1469-7610.2010.02254.x

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*, 91–121. doi:10.1080/10888438.2010.529219

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

Richmond, L. L., Morrison, A. B., Chein, J. M., & Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychology and Aging*, *26*, 813–822. doi:10.1037/a0023631

Spooner, A. L. R., Gathercole, S. E., & Baddeley, A. D. (2006). Does weak reading comprehension reflect an integration deficit? *Journal of Research in Reading*, *29*, 173–193. doi:10.1111/j.1467-9817.2006.00284.x

Swanson, H. L., Kehler, P., & Jerman, O. (2010). Working memory, strategy knowledge, and strategy instruction in children with reading disabilities. *Journal of Learning Disabilities*, *43*, 24–47. doi:10.1177/0022219409338743

Thurstone, T. G., & Thurstone, L. L. (1963). *Primary mental abilities*. Chicago, IL: Science Research Associates.

van den Bos, K. P., Aarnoutse, C., & Brand-Gruwel, S. (1998). Text comprehension strategy instruction with poor readers. *Reading and Writing*, *10*, 471–498. Retrieved from http://link.springer.com/article/10.1023%2FA%3A1007976225000.

van den Broek, P. W. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, *328*, 453–456. doi:10.1126/science.1182594

Yuill, N. M., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge, UK: Cambridge University Press.

Received 5 November 2012; revised version received 11 April 2013