

Working memory and updating processes in reading comprehension

PAOLA PALLADINO, CESARE CORNOLDI,
ROSSANA DE BENI, and FRANCESCA PAZZAGLIA
University of Padua, Padua, Italy

In this study, we examine the relation between reading comprehension ability and success in working memory updating tasks. Groups of poor and good comprehenders, matched for logical reasoning ability, but different in reading comprehension ability, were administered various updating tasks in a series of experiments. In the first experiment, the participants were presented with lists of words, the length of which (4–10 words) was unknown beforehand, and were required to remember the last 4 words in each series. In this task, we found a decrease in performance that was related to longer series and poor reading ability. In the second experiment, we presented lists of nouns referring to items of different sizes, in a task that simulated the selection and updating of relevant information that occurs in the on-line comprehension process. The participants were required to remember a limited, predefined number of the smallest items presented. We found that poor comprehenders not only had a poorer memory, but also made a greater number of intrusion errors. In the third and fourth experiments, memory load (number of items to be selected) and suppression request (number of potentially relevant items) were manipulated within subjects. Increases in both memory load and suppression requests impaired performance. Furthermore, we found that poor comprehenders produced a greater number of intrusion errors, particularly when the suppression request was increased. Finally, in a fifth experiment, a request to specify the size of presented items was introduced. Poor comprehenders were able to select the appropriate items, although their recall was poorer. Altogether, the data show that working memory abilities, based on selecting and updating relevant information and avoiding intrusion errors, are related to reading comprehension.

Several studies have considered the role of working memory—that is, the temporary system that maintains and processes information (Baddeley, 1986)—in reading comprehension. Indeed, it has been shown how the reader must keep in mind information that is useful for interpreting successive parts of the text. For example, in order to solve an anaphora, a reader needs to recover the antecedent of a pronoun (Clark & Sengul, 1979; Daneman & Carpenter, 1980; Ehrlich & Rayner, 1983). Given the limits to working memory capacity, good use of it does not consist in simply maintaining as much information as possible, but in continuously selecting and updating this information. A good comprehender is capable of choosing the most relevant information in the context which he or she is analyzing (Brown, Armbruster, & Baker, 1986; Garner, 1987). He or she is also able to utilize that particular information at the right moment in order to interpret a certain part of the text. On the other hand, the amount of information kept available in the working memory system

can become too broad or even irrelevant. A good comprehender must, therefore, update the available information contained in working memory.

The aim of the present work was to study the relation between the ability to update information contained in working memory and reading comprehension ability. In the literature, this aspect has not been considered, even though particular attention has been paid to the updating processes of mental representations maintained during comprehension. We can distinguish between generally automatic updating in relation to what has been read beforehand (automatically encoded information), which guarantees local coherence, and a somewhat more controlled updating that involves conscious, strategic processes. These processes occur, for example, when it becomes necessary to interpret a portion of the text when the previous representation is inadequate (see, e.g., McKoon & Ratcliff, 1992; van Dijk & Kintsch, 1983). Various researchers (e.g., de Vega, 1995; Glenberg & Langston, 1992; Morrow, Bower, & Greenspan, 1989) have observed that the participant updates the representation of the text in order to understand the new information supplied by the text. In particular, Robertson and Gernsbacher (1996) found that the reader keeps track of the character being discussed in a text and updates the reference when a new character is introduced. The capacity to recall the previous character's name is therefore reduced. The authors have

The present research was partially supported by CNR grants and MURST-40% grants to the second and third authors. We thank Chiara Braga, Barbara Carretti, Martina Di Pieri, Anna Maffezzoni, and Maria Carmela Spoto for their help in data collection. Correspondence concerning this article should be addressed to C. Cornoldi, Dipartimento di Psicologia Generale, Università di Padova, via Venezia 8, 35131 Padova, Italy (e-mail: cornoldi@psico.unipd.it).

interpreted these data, using the structure building framework (Gernsbacher, 1990), as proof of an updating process (a shift to building a new substructure) whereby new information leads to the suppression of previous information.

In this study, we examine the relation between working memory updating and the reading comprehension process from the point of view of individual differences.

The present research used the methodology of comparing matched groups to see whether they differed in their capacity to update working memory. With respect to the well-known memory tests (e.g., Daneman & Carpenter, 1980), a working memory updating test has the advantage of measuring both aspects of working memory implicated in text comprehension—not only the maintenance of selected information, but also its updating.

At first, we considered an updating ability test, developed by Morris and Jones (1990), that follows a procedure presented by Pollack, Johnson, and Knaft (1959). The test requires the participant to listen to a sequence of letters, varying in length (4–10), and afterward to remember only the last 4 elements. Since the length of the series is unknown, it is assumed that each element is initially held in memory and can only be updated as the presentation continues. Morris and Jones consider this a measure of the central executive component of the working memory, which is also hypothesized to be related to reading comprehension ability (Baddeley, 1990).

In five experiments, the present research tested five different groups of people with low text comprehension ability, as compared with participants matched for age, sex, education, and logical ability but with good comprehension ability. In the first experiment, the participants were administered the previously mentioned updating task used by Morris and Jones (1990), in this case replacing the letters with words. In the second experiment, the memory updating task was made more similar to the updating task in the comprehension process. It required participants to select information, with reference not to a simple ordinal position criterion (such as the last four words), but to a predetermined semantic criterion (the words that mainly satisfied the criterion; in our case, the words referred to the smallest items). In the third experiment, we wanted to generalize the results to a sample of younger participants and to evaluate the role of required memory load and quantity of potentially relevant information. In the fourth experiment, we analyzed the typology of errors made in the updating task. In the fifth experiment, we controlled for the possibility that poor comprehenders' difficulty in working memory updating tasks was not due simply to a failure to identify the items that had to be remembered.

In summary, the aim of the five experiments was to test the hypothesis that comprehension ability is related to the ability to update relevant information maintained in working memory. This hypothesis is the consequence of a model that assumes that, among the operations required in the comprehension process, there are operations that involve the selection and use of information that is relevant to interpreting the passage. Similarly, in the work-

ing memory updating test, when an individual is invited to remember, he or she must give the information that is relevant at that time. In order to do this, he or she must have updated the relevance and, therefore, the degree of accessibility of information processed, thereby reducing the accessibility of information that is (certainly or probably) no longer useful. Success in the task is linked to remembering the pertinent information and also to controlling for nontarget information with respect to items that were potentially relevant but did not satisfy the selection criterion. This information, if not updated, can disturb the memory of pertinent information but can also be erroneously remembered, thus producing an intrusion error. It has recently been observed (De Beni, Palladino, Pazzaglia, & Cornoldi, 1998) that intrusion errors in working memory tasks are highly related not only to success in the working memory tasks themselves, but also to reading comprehension ability.

EXPERIMENTS 1–5

Method for the Selection of Participants

For all five experiments, we selected a group of *good comprehenders* and a group of *poor comprehenders* according to the following criteria. In a reading comprehension test, good comprehenders were required to have a performance at least 1 *SD* higher than the average score expected for their age, and the group of poor comprehenders a performance at least 1 *SD* lower. The reading comprehension test was based on the MT battery (Cornoldi, Rizzo, & Pra Baldi, 1991), the only standardized test for assessing reading comprehension available in Italy. The test includes passages followed by multiple choice questions. Groups were matched for age and general logical intelligence, which was examined with a battery of logical reasoning subtests based on items (numbers and spatial patterns) present in standard group intelligence tests (Anastasi, 1988).

EXPERIMENT 1

The main goal of Experiment 1 was to test the hypothesis that poor comprehenders' performance is lower than that of good comprehenders in a working memory updating task. The participants were required to update not just letters, as in the original updating task (Morris & Jones, 1990), but more meaningful and semantically relevant items—that is, concrete, highly familiar nouns. Furthermore, they were tested with a word span test in order to exclude the possibility that failure in the updating test was due to poor short-term memory.

Method

Participants. Sixteen poor comprehenders (4 males and 12 females), between 19 and 21 years of age, all undergraduate students at the University of Padua, and 14 good comprehenders (4 males and 10 females), with similar education, age, short-term memory, and logical intelligence, agreed to participate in the experiment.

Short-term memory was measured by a forward word span test. Procedures were drawn from the Wechsler Intelligence Scale–Revised (Wechsler, 1987), standardized for the Italian population (Orsini & Laicardi, 1997). The test, which was administered individually, assesses serial recall of strings of words that are presented orally. Different from the original procedure (Wechsler, 1987), the participants were presented with only one list for each length. The

Table 1
Means and Standard Deviations of Number of Words
Recalled in the Updating Task in Experiment 1

Number of Words	Words Recalled	
	<i>M</i>	<i>SD</i>
4 (no updates)	14.90	1.95
6 (two updates)	13.2	2.46
8 (four updates)	12.63	2.14
10 (six updates)	11.47	2.11

score was based on the greatest list length a participant was able to recall completely in the correct order (span). The mean span was 4.94 ($SD = 1.44$) for poor comprehenders and 5.14 ($SD = 1.10$) for good comprehenders, with a slight nonsignificant difference between the groups ($t = 0.043$, $p = .67$).

Materials and Procedure. Sixteen lists of words were presented. The words were bisyllabic, highly familiar nouns taken from the Cornoldi (1974) repertoire. The 16 lists were divided into four groups, varying by number of words (4, 6, 8, or 10). An example of an 8-word list was the following: *forno* (oven), *latte* (milk), *duomo* (cathedral), *fiato* (breath), *vasca* (bath), *zucca* (pumpkin), *succo* (juice), *cassa* (box), and *baffo* (mustache). The lists were presented orally, in a fixed order, using a tape recorder and balancing the presentation order of lists of different lengths. Within each list, the words were presented at the rate of 1 word per second. The end of the list was signaled by a beep, and the participant was immediately required to write the last 4 words in the list. If he or she could not recall a word, it had to be omitted (in contrast to Morris & Jones's [1990] procedure, the participants were not required to remember four items). The participants were tested in small groups of 3 or 4. The test was preceded by a warm-up session with two lists of each length. The experiment took approximately 15 min.

Results

Table 1 presents the mean numbers of correctly reproduced words for the different length lists. A word was considered correctly reproduced only if it was recalled in the correct position. When a list was only partially recalled, a recalled word was considered ordered if it was not presented before the word just recalled. A 2×4 analysis of variance (ANOVA) for a mixed design revealed significant effects of group [$F(1,28) = 4.98$, $MS_e = 7.86$, $p = .034$; good comprehenders, $M = 54.64$, $SD = 5.8$; poor comprehenders: $M = 50.06$, $SD = 5.43$] and of list length [$F(3,84) = 18.27$, $MS_e = 3.35$, $p < .001$]. We also computed the errors, but the two groups did not differ significantly in the mean number of either intrusion errors ($M = 3.31$, $SD = 1.70$ for the poor comprehenders; $M = 2.64$, $SD = 2.56$, for the good comprehenders; $p = .40$) or order errors ($M = 1.50$, $SD = 1.41$, for the poor comprehenders vs. $M = 1.00$, $SD = 1.36$, for the good comprehenders; $p = .33$). Overall, the number of responses of both good and poor comprehenders (58.28 and 54.87, respectively) was below the maximum possible value of 64.

Discussion

The results of Experiment 1 confirm that even when poor comprehenders are matched with good comprehenders in a series of abilities, including short-term memory,

they may present a deficit in working memory tasks involving central control (see Daneman & Merikle, 1996; De Beni et al., 1998). In other words, the control functions of working memory seem related to text comprehension ability. This relation was extended to a working memory test that was not originally devised for examining the relation between working memory and comprehension. The updating test we used is considered a measure of the central executive component of working memory (Morris & Jones, 1990; Van der Linden, Bredart, & Beerten, 1994). The fact that poor comprehenders had a lower performance than good comprehenders in this task, but not in a standard span task, confirms that, within Baddeley's (1990) tripartite working memory model, the comprehension process engages the central executive more than it does the articulatory loop. It should be noted that we found a significant effect of list length, but no interaction between groups and list length. These data demonstrate that an increase in list length produces a more complex request, thus impairing performance. However, although this aspect is important in principle, it is not necessarily crucial to reading comprehension ability. The two reading comprehension ability groups were similarly affected by an increase in list length. In particular, it should be noted that poor comprehenders also performed worse with lists of four items in which no substitution–updating was apparently required. This result suggests that memory updating does not involve only simple maintenance and substitution, but probably more general control and coordination of information activation/suppression in working memory.

The results also offer an insight into the role of updating ability in reading comprehension. In fact, both the updating and the reading comprehension tasks require the participant to continuously update information and, when required, find appropriate information. This requirement cannot be one-way only but, rather, takes place through a flexible and strategic use of working memory. In fact, only a few participants reported that they were simply maintaining the last four words, continuously substituting the first ones. As the presentation rate was very rapid and the quantity of material was unknown and could have been small, many participants reported that they had partially selected and updated the material during the presentation and partially completed the updating operation when required to recall the words.

EXPERIMENT 2

The working memory updating task used in Experiment 1 had the advantage of requiring a continuous process of refreshing relevant information with semantically significant items, but it had the disadvantage of simulating rather superficially the updating processes involved in real comprehension tasks. In fact, reading comprehension requires the reader to select and update on the basis of the semantic relevance of information, not just on the basis of information position. For this reason, we developed a task in which updating had to be made on the basis

of a semantic criterion (item size), rather than simply by using a presentation order criterion. In this version of the task, the length of the list was always the same, but the number of relevant elements varied within the list, as did the number of elements that had to be remembered. We also introduced a final recall test in order to test the hypothesis that the memory deficit of poor comprehenders is not general, but specifically involves target items.

Method

Participants. Eighteen poor comprehenders and 20 good comprehenders, with characteristics similar to those of the participant groups in Experiment 1, participated in this experiment.

Materials and Procedure. Twenty-four lists of 12 words were presented. Each list contained familiar nouns that referred to either objects/animals measurable by size (relevant information; in a number varying from 5 to 10) or abstract entities (filler items). Nouns of sizable objects/animals were selected in a pilot study. In the pilot, five judges rated the discriminability of the item sizes in the lists. From the ratings, it appeared that the sizes of the items in each list were immediately evident and clearly different from each other.

The instructions emphasized that the participant would be presented with lists including filler abstract items and concrete items, the size of which had to be considered. Only for concrete items did he or she have to select the smallest. For half the lists, the participants had to remember the three smallest items, and for the other half, the five smallest. The words were presented orally using a tape recorder, at the rate of one word per second. Each participant was tested individually and gave a verbal response after each list. An example of a list was *incontro* (meeting), *senso* (sense), *picchio* (woodpecker), *passione* (passion), *legge* (law), *mucca* (cow), *felicità* (happiness), *quantità* (amount), *bruco* (caterpillar), *agnello* (lamb), *festa* (feast), and *rana* (frog), and the three words to be remembered were *woodpecker*, *caterpillar*, and *frog*. At the end of the experiment, the participant was invited unexpectedly to remember all the words which had been presented. The recall test was written, and the participant had 7 min to complete the test.

Results

Table 2 presents the mean numbers of words recalled correctly for the two groups, their overall intrusion errors (nouns presented in the list, but not the smallest ones), and the mean numbers of words that were recalled in the final test divided by category. The poor comprehenders gave a significantly lower performance in the updating task, recalling 80.16 words (out of a maximum of 96), versus a mean recall of 85.95 for good comprehenders, a highly significant difference [$t(36) = 3.21, p = .003$]. The difference in intrusion errors was significant, being nearly double for the poor comprehenders [$t(36) = 2.74, p < .01$].

Regarding final recall (Table 2, bottom), the only significant difference between the groups concerned the target words [$t(36) = 2.14, p = .039$], a result that substantially reflects the good comprehenders' better working memory recall of target items. In the other final recall measures, the two groups gave a similar performance, which suggests that in other respects, the memory of the two participant groups was similar. Overall, the percentage of sizable, but nontarget, items (i.e., the larger items) recalled (18.63%) was lower than the percentage of target

Table 2
Means and Standard Deviations for Poor and Good Comprehenders in Indices of Updating and Final Recall Tasks (Experiment 2)

Task	Poor Comprehenders		Good Comprehenders	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Updating				
Words recalled	80.16	6.28	85.95	4.79
Intrusions	10.67	5.49	6.35	4.19
Unexpected final recall				
Target words	26.05	5.52	30.30	6.57
Relevant items	14.55	5.04	16.75	6.46
Filler items	3.39	3.07	2.50	2.15

items (30.76%), but much greater than the percentage of filler items [2.73%; $F(2,72) = 261.33, MS_e = 28.66, p < .001$].

Discussion

The present experiment confirms that poor comprehenders have difficulty with verbal working memory tests and with updating tasks in particular. In this case, in contrast to Experiment 1, updating had to be done on the basis of a semantic criterion. The updating process therefore reflected the updating operations more directly involved in the comprehension, since text comprehension requires that relevant information is extracted and made highly accessible in working memory. The pool of accessible information is then updated in favor of new, more relevant, incoming information. It must be noted that poor comprehenders recalled a lower number of target items in both the immediate and the final tests, suggesting that appropriate selection and updating of information had long-term effects.

In the present case, we also found a difference between groups in the mean number of intrusion errors. Poor comprehenders made more intrusions when the intrusive information initially had to be processed deeply—that is, for concrete (sizable) nouns. The results of this experiment replicate De Beni et al.'s (1998) observations, which showed that poor comprehenders make more intrusion errors than good comprehenders in listening span tests and that these errors are more frequent for more extensively processed information. However, differences in listening and updating task procedures do not allow us to attribute the effects to exactly the same factors. Those participants interviewed after the experiment reported that they maintained many of the relevant items activated to a different extent. Intrusion errors could have been due to memory noise, wrong levels of activation, inappropriate selection when the response was required, or other factors. In particular, the participants' difficulty could have been due to an increase either in the number of relevant items (leading to an increase in the request to suppress nontarget items) or in the memory request (leading to an increase in the number of items to be recalled). A further experiment tested between these possible explanations more directly.

EXPERIMENT 3

In Experiment 3, we examined the consequences of different memory and suppression requests in the updating task administered in Experiment 2. The task was presented to new groups of good and poor comprehenders. Furthermore, the larger size of the initial available sample made it possible to select a larger group of poor comprehenders than in the other experiments. In this experiment, we tested the hypothesis that updating is affected by increases in suppression and memory requests. The suppression request was considered higher when the participant had to extract target information from a larger pool of relevant information. Consequently, more information must be activated, examined, and temporarily maintained in memory, and in turn, more information must be suppressed. The memory request was considered higher when the participant had to remember a larger number of target items. Consequently, he or she had to hold in memory and/or extract a greater number of elements, with an increasing amount of activity required. Once again, we also examined the effects of memory and suppression requests in a working memory updating task (performance and errors) for two groups that differed in reading comprehension ability. We hypothesized that higher memory and suppression requests would tend to overload the working memory system of both groups, impairing performance. Furthermore, if poor comprehenders have difficulty maintaining relevant information and rejecting what has become irrelevant, they should also have a poorer performance and present more intrusion errors with a higher suppression request.

Method

Participants. Thirty poor and 30 good comprehenders, between 11 and 15 years of age, attending a junior high school in Lodi, a town in Northern Italy, participated in this experiment.

Materials and Procedure. As in the preceding experiment, 24 lists of 12 words were presented. The words in the lists were mostly the same as those used in the preceding experiment; the new words had similar characteristics. Relevant information was made up of animal and object nouns distributed in the different lists, balanced within-subjects. The 24 lists were divided into four categories, according to the number of relevant items (high and low suppression required) and the number of items to be recalled (high and low memory request in loading and retrieval operations). In the low-suppression lists there were two irrelevant items, whereas there were five in the high-suppression lists. The participant had to remember three items in the case of low memory load and five items for high memory load. Lists for each category were presented in a block. The following is a list with high suppression and memory requests: *penna* (pen), *relazione* (relationship), *portafogli* (wallet), *culla* (cot), *scala* (ladder), *occhiali* (glasses), *dolore* (pain), *quadernone* (exercise book), *gomma* (rubber), *treno* (train), *sedia* (chair), and *violino* (violin). The participants were required to remember the five smallest items: pen, wallet, glasses, exercise book, and rubber. The order of the blocks of lists was balanced between subjects. There were four practice trials for each category. In order to avoid recency effects, at the end of the working memory test, the participants were asked to do an interpolated task, consisting of counting backward by two from 20 to 0 before doing an unexpected written free recall test of all the items that had been presented. Instructions,

practice, modality of item presentation, and the items themselves were the same as those in Experiment 2. The participants were individually tested in a quiet room in their school.

Results

Figure 1 presents the mean percentages of words from the four categories of the lists that were remembered by the two groups (percentages, rather than absolute values, were considered because the sets of items to be recalled varied according to the category of lists). A $2 \times 2 \times 2$ mixed design ANOVA (group \times suppression request \times memory request) revealed a significant group effect [$F(1,58) = 21.32$, $MS_e = 430.63$, $p < .001$], a significant effect of suppression request [$F(1,58) = 20.48$, $MS_e = 91.64$, $p < .001$], and a significant effect of memory request [$F(1,58) = 343.72$, $MS_e = 73.23$, $p < .001$]. Poor comprehenders gave a poorer performance, and increased suppression and memory requests impaired recall.

Figure 2 presents the mean numbers of intrusion errors for the two groups divided according to list categories. A $2 \times 2 \times 2$ ANOVA revealed that all the main effects were significant [groups, $F(1,58) = 6.50$, $MS_e = 9.54$, $p = .013$; suppression request, $F(1,58) = 15.15$, $MS_e = 2.54$, $p < .001$; memory request, $F(1,58) = 15.76$, $MS_e = 1.78$, $p < .001$]. Furthermore, the group \times suppression request interaction was significant [$F(1,58) = 5.15$, $MS_e = 2.54$, $p = .027$]. From Figure 2, it can be seen that increases in suppression and memory request produced a higher number of intrusions and that the interaction was due to the fact that poor comprehenders were particularly disturbed by the high-suppression lists. A closer look at the number of intrusions produced by the poor comprehender group revealed a high level of variability. This could be due to the fact that poor comprehenders used a greater variety of strategies than did good comprehenders.

The number of recalled filler items was particularly low in the final recall test, so we only considered the number of relevant items recalled, divided according to whether they were a target or had to be suppressed. A 2×2 ANOVA revealed that good comprehenders remembered a significantly higher number of relevant nouns than did poor comprehenders [$F(1,58) = 12.91$, $MS_e = 25.70$, $p < .001$] and that target items were remembered better than nontarget items [$F(1,58) = 33.51$, $MS_e = 21.00$, $p < .001$]. The absence of an interaction was due to the fact that good comprehenders remembered not only the target items better ($M = 18.92$, $SD = 6.37$, vs. $M = 15.21$, $SD = 4.93$), but also the relevant nontarget items ($M = 13.69$, $SD = 4.17$, vs. $M = 10.75$, $SD = 3.34$).

Discussion

The results of Experiment 3, on the whole, replicated the results of Experiment 2 with a different group of participants. It appears that poor comprehenders have difficulty in the updating task, producing both a lower recall and a greater number of intrusions. Furthermore, success in the task is related to the operations of suppressing relevant information and storing and retrieving target infor-

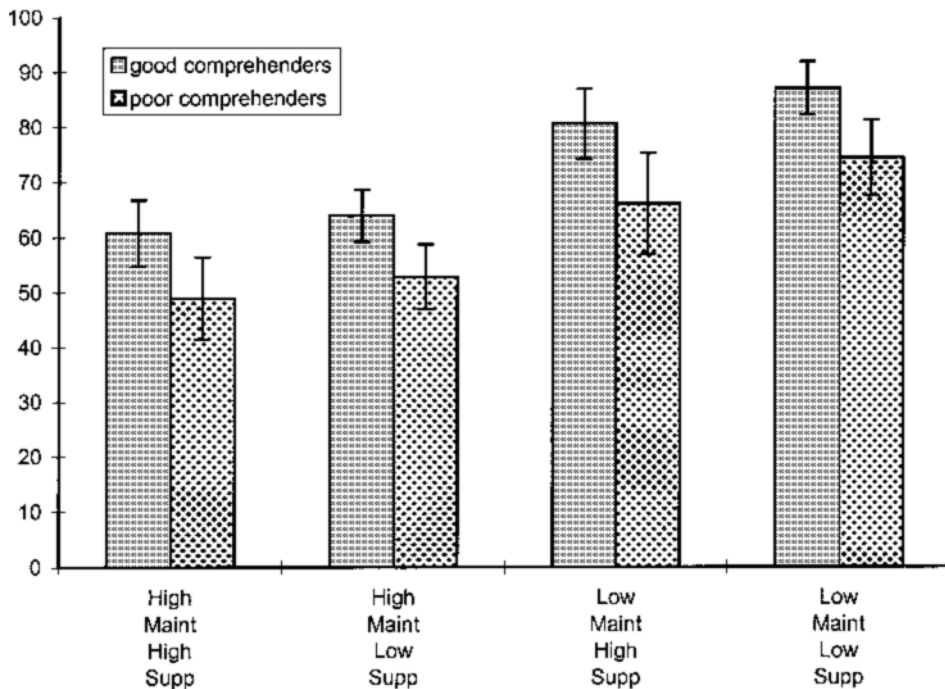


Figure 1. Mean percentages of words correctly recalled by good and poor comprehenders in the Experiment 3 updating task under high or low maintenance (Maint) and suppression (Supp) requests.

mation. Memory and suppression requests affected both recall and intrusion errors, in both groups. However, an interaction between groups and suppression requests was observed for intrusion errors. This result suggests that poor comprehenders have a specific difficulty suppressing relevant information when there is a high number of items to be suppressed. In fact, good comprehenders made a similar number of intrusion errors with high- and low-suppression lists. On the contrary, poor comprehenders' intrusion errors in high-suppression lists were twice as many (6.27 vs. 3.73), as compared with the low-suppression lists. This effect was specific to intrusions. By contrast, the two groups appeared to be similarly affected by increased memory request (see Figure 2). These data suggest that poor comprehenders' difficulty in avoiding intrusion errors is not due to an increased memory load or retrieval effort but is related to the quantity of relevant information that must be suppressed. An increase in the number of items to be selected affects the two groups similarly; an increase in items to be suppressed specifically adds to the difficulty poor comprehenders have in avoiding intrusion errors.

The data on final recall again showed that good comprehenders had a better memory for target items, which substantially mirrored the difference already found in working memory recall, as in the preceding experiment. However, in this case, we also found an unexpected difference between groups in the recall of relevant nontarget information. It should be noted that the performance pattern of the participants in this experiment was gener-

ally different from the pattern we found in the preceding experiment, since the absolute values of target items recalled were lower in this experiment than in the preceding one. It is possible that the interpolated task introduced in the present experiment had the effect of impairing recall of items and increasing the role of long-term semantic abilities that may be superior in good comprehenders (Cornoldi & Oakhill, 1991).

EXPERIMENT 4

The aim of Experiment 4 was to obtain further evidence concerning the effects found in the preceding experiments and to explore the characteristics of intrusion errors. In fact, an intrusion error could be due to an item either in a preceding list or in the same list. Intrusions of items in the same list may be considered as working memory errors, because the intrusive elements are presumably maintained in working memory. Poor comprehenders, who have a poorer working memory system, should be differentiated by this latter measure, but not by the former. Furthermore, De Beni and Palladino (2001) have argued that it may be possible to distinguish between two types of intrusions: intrusions of relevant items that can be immediately excluded from the pool of target items (*immediate intrusions*) and those of relevant items that can only later be excluded from the set of target items (*delayed intrusions*). Let us consider the list introduced previously as an example. The items were presented in the following order: pen, relationship, wallet, cot,

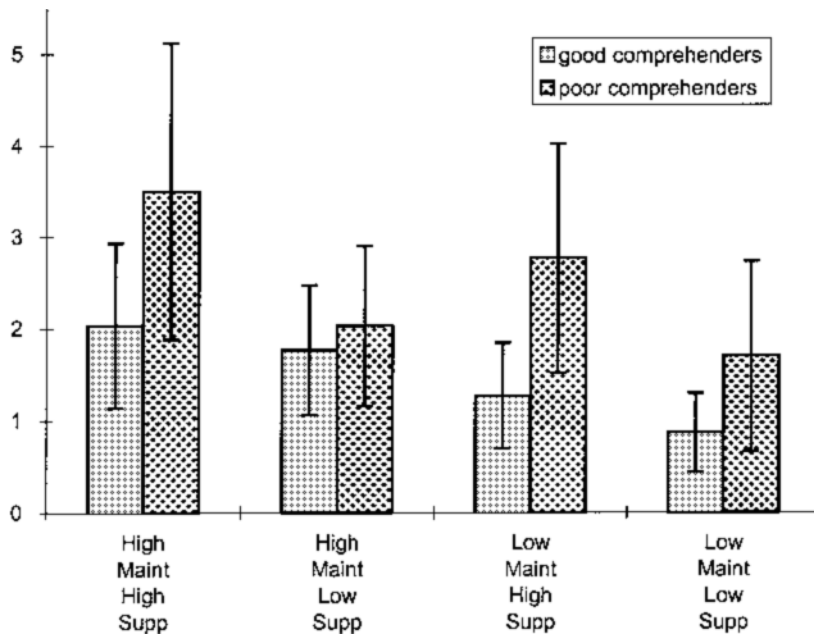


Figure 2. Mean numbers of intrusions in the Experiment 3 updating task made by good and poor comprehenders under high or low maintenance (Maint) and suppression (Supp) requests.

ladder, glasses, pain, exercise book, rubber, train, chair, and violin, where pen, wallet, eyeglasses, exercise book, and rubber were the target items. The other relevant items, cot and ladder, had to be considered for a while within the pool of relevant items and then excluded (delayed exclusions). Although larger than the target items, cot and ladder were presented when the working memory was not yet occupied by smaller items. By contrast, the items train, chair, and violin could immediately be considered as being nonappropriate (immediate exclusions), because they were presented at the end of the list when smaller items were already in memory.

The other type of intrusion concerned the items in preceding lists that we assume no longer pertain to working memory but, rather, to long-term memory. May, Hasher, and Kane (1999), in a study with older participants, found that they were more affected by proactive interference produced by items from preceding lists. If intrusion errors generically involve all the memory systems and, therefore, all kinds of information, the distinction between items belonging to the same and those belonging to preceding lists should not have any influence. If, on the contrary, the poor comprehenders' intrusion problem is related to the manipulation of highly accessible information, intrusions would be specifically due to items from the latest list, but not those from preceding lists.

Method

Participants. Thirteen poor comprehenders and 17 good comprehenders, similar in characteristics to the groups of participants tested in Experiment 1, participated in the experiment.

Materials and Procedure. The materials and procedure were the same as those in Experiment 3. In this case, final recall was not required.

Results

Table 3 presents the mean percentages of nouns contained in the four types of lists that were recalled by the two groups. A $2 \times 2 \times 2$ ANOVA revealed significant effects of group [$F(1,28) = 6.20, MS_e = 208.70, p = .019$] and memory request [$F(1,28) = 59.60, MS_e = 87.76, p < .001$]. The effect of suppression request only approached significance [$F(1,28) = 3.80, MS_e = 82.48, p = .061$].

In the analysis of intrusion errors, we considered a source of error factor by considering intrusions from the same list, as well as those from preceding lists (see Table 4). A 2×2 ANOVA (groups \times source of error) showed a main significant effect of type of intrusion [$F(1,28) = 85.89, MS_e = 20.70, p < .001$], owing to the fact that intrusions were mainly items from the same list, and a significant interaction between groups and type of intrusion [$F(1,28) = 4.93, MS_e = 20.70, p = .035$]. Single-cell comparisons revealed that poor comprehenders had significantly more intrusions of items in the same list [$t(28) = 2.13, p = .042$]. On the contrary, the two groups were not significantly different for intrusions from items in preceding lists, and good comprehenders made an even higher number of errors.

For intrusions from items in the same list, we considered further the immediate intrusions (sizable objects that, owing to their size, can be immediately excluded) and the delayed intrusions (items that can be excluded only

Table 3
Mean Percentages and Standard Deviations
of Correctly Recalled Words for Poor
and Good Comprehenders (Experiment 4)

Type of List	Poor Comprehenders		Good Comprehenders	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
High load, high suppression	67.69	13.08	68.82	9.71
High load, low suppression	70.51	9.31	78.04	8.17
Low load, high suppression	80.77	11.70	87.91	10.25
Low load, low suppression	79.49	14.94	90.20	7.99

after a while). This contrast must be considered with caution, since the items could not be balanced. We did not find a greater increase in delayed intrusions in the poor comprehender group; in fact, poor comprehenders made 37% more delayed intrusions, a proportion that was even lower than that observed for immediate intrusions (61% increase). However, in both groups, delayed intrusions were more frequent than immediate intrusions.

Discussion

The results of Experiment 4 substantially replicated the results of the preceding experiments. Poor comprehenders had poorer working memory, as is manifested by their performance in an updating memory task. In the present case, the differences between groups were less dramatic, since the poor comprehenders were selected on the basis of a less strict criterion.

The main goal of this experiment concerned the possibility of finding different intrusion error patterns in the groups of poor and good comprehenders. In fact, we found that poor comprehenders made more intrusion errors only with respect to items in the latest list and independent of the exclusion delay within the list. This result confirms the hypothesis that poor comprehenders' intrusion errors are therefore specifically related to the working memory system. Since the critical variable differentiating poor and good comprehenders is comprehension ability, the results suggest that a similar pattern of intrusion errors is also related to text comprehension. The effect seems to be due to high memory strength information, rather than to a weakened trace and loss of memory discriminability, as was found by De Beni et al. (1998).

The present results suggest that task difficulty in the present study was not related to time of permanence in memory before exclusion but, rather, to a more complex series of strategies that may include not only the possibility of immediate exclusion, but also the possibility of maintaining a large part of the material at different levels of activation.

The absence of a difference between groups in number of intrusions from a previous list does not necessarily mean that the two groups (poor and good comprehenders) were affected to a similar extent by proactive interference. It has been shown that only a small part of the long term memory interference effect is emphasized by an increase in the number of intrusions (Barnes & Underwood, 1959). However, this result suggests that long-term mem-

ory interference is less crucial than working memory interference in the difference between groups in the updating task.

EXPERIMENT 5

The data of the last three experiments were coherent in showing that poor comprehenders have difficulty updating information in working memory. The task we devised to test this assumption—that is, the request to remember only the smallest items—was very simple and easy to accomplish not only for young adults, but also for children. However, it is possible that the same task became difficult in the context of the updating task. In other words, it is possible that poor comprehenders, having difficulty in the updating task, became disoriented even with the simple task of evaluating item size. In this case, low performance in memory for the smallest items would not be due to a lack of updating, but to a failure in recognition of the smallest items owing to attention lapse in the complex double task. In order to test this possibility, we ran a further experiment that explicitly required the participants to indicate the size of each item presented. If poor comprehenders have difficulty in this evaluation, it may be an indication that their poorer performance in the updating task is related to an inability to correctly identify the target items, rather than to updating per se.

It must be noted that this new request emphasizes the importance of the smallest items, and this manipulation could also affect overall performance in the updating task. In fact, we assumed that poor comprehenders' difficulty in updating was due to the fact that they were also paying attention to items that were irrelevant with respect to the critical feature of *being small*. Here, this inappropriate focusing should be reduced, with the probable consequence of a reduction in poor comprehenders' difficulty.

Method

Participants. Fourteen poor comprehenders and 15 good comprehenders similar in characteristics to the groups tested in Experiments 1, 2, and 4 were selected.

Materials and Procedure. The materials were the same as those in Experiments 3 and 4, and the procedure was the same as that in Experiment 4, except that the participant had to give a size estimation for each item. After presentation of each item, the participant had to indicate, for the sized items, its size on a scale from 0 to 100,

Table 4
Means and Standard Deviations of Different
Types of Intrusion Errors Produced by Poor
and Good Comprehenders (Experiment 4)

Intrusions	Poor Comprehenders		Good Comprehenders	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Immediate	5.38	3.73	3.35	2.21
Delayed	9.77	5.13	7.12	3.30
Total from the same list (immediate + delayed)	15.15	7.58	10.47	4.37
Total from preceding lists	1.54	1.66	2.12	1.11

Table 5
Mean Numbers of Correctly Individuated Items, Mean Percentages
of Correctly Recalled Words in the Different Lists, and Mean Numbers
of Intrusion Errors by Poor and Good Comprehenders (Experiment 5)

	Poor Comprehenders		Good Comprehenders	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Correct identifications (max = 96)	93.71	3.32	94.00	1.89
High load, high suppression (%)	66.19	8.56	71.11	9.97
High load, low suppression (%)	65.48	9.92	75.56	8.03
Low load, high suppression (%)	82.14	13.64	85.19	7.76
Low load, low suppression (%)	85.32	9.65	89.63	6.25
Intrusions	4.57	4.35	5.13	3.36

whereas for the filler items he or she just had to say “unsized.” The participants were instructed on how to do the task and were presented with examples of items of different sizes in order to become familiar with the 0–100 scale. They were also instructed to give an immediate response, because the presentation rate could not be changed. At the end of the list, the task proceeded in the same way as in the preceding experiments.

Results

The participants did not have difficulty with the task, since they were able to give immediate responses for the size request. The fact that neither the good nor the poor comprehenders had difficulty with the size evaluation is confirmed by the very high numbers of smallest items that actually received the *smallest* evaluations. This occurred indifferently for all the types of lists, including the most difficult ones. Table 5 presents the main results of the study.

The number of smallest items correctly identified was very close to the maximum number of 96 in both groups, and the difference between groups was not significant ($p = .78$). However, poor comprehenders still had difficulty in the updating task—that is, in recalling the smallest items. A $2 \times 2 \times 2$ mixed design ANOVA (group \times suppression request \times memory request) on the percentages of correct responses revealed a significant group effect [$F(1,27) = 4.39, MS_e = 206.19, p < .05$]. Furthermore, the main effects related to the nature of the lists were also significant for both suppression request [$F(1,27) = 5.27, MS_e = 44.27, p < .05$] and memory request [$F(1,27) = 135.53, MS_e = 54.62, p < .001$].

Despite a difference in correct recalls, the two groups did not differ in the number of intrusion errors ($p = .70$). In fact, the new task requirement had a general effect of reducing the tendency to recall large objects. Intrusion errors in both groups were lower than those in the preceding experiments, with reductions even greater than 50%.

The last line of Table 5 presents the mean numbers of intrusion errors owing to the recall of items from the same list that were not the smallest ones. The data concern immediate and delayed suppressions considered together, since the pattern of errors was similar for the two. Poor comprehenders had a slight, but not significant, tendency to make more immediate intrusions (2.43 vs. 2.14), and good comprehenders a tendency to make fewer immediate intrusions (1.80 vs. 3.33). The importance of immediate

intrusions in the task is confirmed by the significant correlation between the number of immediate intrusions and the number of correct recalls (Pearson's $r = -.56, p < .01$) found when considering all the participants, whereas the corresponding correlation between delayed intrusions and correct recalls did not reach significance ($r = -.04$).

Discussion

The results of Experiment 5 clearly show that the poor comprehenders' difficulty in the updating task proposed in the preceding experiments was not due to an inability to evaluate the size of items and then to identify the smallest ones. Despite the accuracy of their ratings, the poor comprehenders had a lower recall of the target (i.e., smallest) items. Therefore, the difficulty of the poor comprehenders was not due to inappropriate identification but seems to have been due to an inability to select the appropriate items during recall. However, in this case, the poor comprehenders' working memory difficulty was not related to a general increase in the number of intrusion errors but, possibly, to an inability to immediately suppress large items (immediate intrusions). The two groups made a very low number of intrusions. This dramatic decrease appears to have been due to the request to evaluate the size of the items, which presumably caused the participants' attention to be focused on the items that received absolute low ratings, reinforcing the appropriate selection. In fact, many of the participants reported that sometimes during recall they were able to remember the numbers given in the ratings even better than the related items. This occurred in particular for the items that received the lowest ratings. For example, a participant could say “Well, I remember there was an item I gave a rating of 5 to, now I will try to retrieve it.” In our opinion, this focus on the smallest items contributed further to reducing the focus on the larger items, thus reducing the probability that the latter would be remembered.

GENERAL DISCUSSION

A large body of literature (for a meta-analysis, see Daneman & Merickle, 1996) has shown that working memory is associated with reading comprehension ability. In fact, reading comprehension requires that parts of the text are temporarily held in a working memory system to be ana-

lyzed and integrated. A critical working memory operation involved in reading seems to be updating relevant information. However, no previous research has shown any relation between reading ability and the ability to carry out a working memory updating task. The present study offers considerable evidence in favor of this relation. In fact, five different groups of poor comprehenders presented a similar pattern of difficulty in updating tasks.

Experiments 2, 3, and 4 also show that poor comprehenders, under certain conditions, tend to make more intrusion errors—that is, they remember more items that are relevant but not target items. This last result is intriguing, because it adds evidence to the hypothesis that success in working memory and comprehension tasks is related not only to the ability to recall selected information, but also to the ability to suppress other potentially relevant but not target information (e.g., De Beni et al., 1998; Gernsbacher, Varner, & Faust, 1990). The fact that the same effect was not found in Experiments 1 and 5 confirms that the two updating tasks did not measure exactly the same process. Experiments 1 and 5 highlighted, in particular, the target items through either final position (Experiment 1) or the request to explicitly evaluate the size of the objects. Since the intrusion error appears to be affected by the stress given to nontarget items (De Beni et al., 1998; Passolunghi, Cornoldi, & De Liberto, 1999), appropriate emphasis given to target items should have the consequence of reducing the emphasis on, and tendency to recall, nontarget items.

Furthermore, we found there to be general effects owing to poorer performance in the updating task with increases in suppression and memory requests. These data offer a first insight into the nature of the updating process. The memory request effect shows that updating is more complex when more information must be selected, probably because of limitations in the working memory system, which becomes overloaded by the memory load in addition to the operations required by the task. The suppression request effect shows that success in the updating task is related not only to the memory request—that is, the number of elements that must be remembered—but also to the number of elements that must be excluded. Obviously, these elements were also potential targets, and a higher suppression request also implies an increase in memory load, which cannot easily be dissociated from the suppression request per se.

Our observations apparently do not agree with those of Morris and Jones (1990), who found that the number of updates and a concurrent task do not depress updating performance. However, it should be noted that, as well as involving differences in materials and procedure, Morris and Jones's concurrent task concerned the articulatory loop, rather than the more active central components of working memory, as did our suppression and memory request manipulations.

In order to understand completely the psychological mechanisms involved in updating, new and further evidence is necessary. A simple view of updating (Morris &

Jones, 1990) assumes that the participant holds the target items in working memory, and then, when new target items arrive, those that are no longer targets are dropped. For example, it could be assumed that in the updating tasks of Experiments 2, 3, 4, and 5, the participant selects the three (or five) smallest items, and if another small item was presented, it was added to his or her memory, and the largest one maintained was then dropped. However, our data suggest that the process is more complicated. For example, we found that in Experiment 1, although they had a span as high as good comprehenders, poor comprehenders were poorer also with the four-item list updating task—that is, when no drop–substitution process was required [good comprehenders: $M = 15.64$, $SD = 1.08$; poor comprehenders: $M = 14.25$, $SD = 2.32$; $t(28) = 2.15$, $p = .043$]. This result is against the simple view of updating as an “all or nothing” process of maintenance–substitution. In Experiments 4 and 5, we found that poor comprehenders' difficulty was not related more to delayed exclusions. This suggests that their difficulties lie not only in substitution, but also in selection/maintenance. It must also be noted that, if reduction of the activation/elimination from memory of items no longer included in the target pool is immediate, it would be difficult to explain why good comprehenders, who have better updating ability, should also have a better long-term memory for all the items presented.

Altogether, these findings suggest that updating is not a simple inclusion/exclusion process. It is probably a more complex process that attributes different levels of activation to the items presented and continuously updates that level while maintaining a larger set of elements activated. In reading comprehension, the reader should keep various pieces of information available in a similar way, tuning activation level to their relevance and importance, until a final interpretation can be made and a clear text representation or mental model built. Obviously, this theoretical perspective is purely speculative (even if consistent with some literature and reading comprehension models; see, e.g., Gernsbacher, 1990; van Dijk & Kintsch, 1983) and requires experimental evidence.

In explaining why poor comprehenders fail our updating tasks, we must consider the hypotheses advanced for the similar failure of poor comprehenders in other working memory tasks. The ability to process sentences has not been ruled out as a factor explaining the relation of the listening span test to reading comprehension (Daneman & Carpenter, 1980, 1983). However, this explanation can be ruled out, because no sentence processing was required in these updating tasks. Similarly, we can reject the hypothesis that the relation between reading comprehension and updating ability is simply due to very general intellectual ability, since the participants were matched for at least some intelligence components. In Experiment 1, the participants were also matched for their short-term memory ability, confirming that working memory differences are not necessarily related to immediate memory differences (see De Beni et al., 1998).

Engle and coauthors (e.g., Conway, Tuholski, Shisler, & Engle, 1999; Engle, Cantor, & Carullo, 1992) have argued that success in remembering relevant, and in suppressing irrelevant, information in working memory tasks is related to the quantity of cognitive resources available to the working memory system. This perspective seems attractive, since it is compatible with the view that the updating task requires complex management of the activation processes, modulating at each moment the status of all the items. It also takes into consideration the suppression request involved in this task.

The particular results of Experiments 1 and 5 (difference between groups in recall, but not in intrusions) show that, in our task, the participants did not adopt the strategy of using any item at all to make up the expected number of items to be recalled. The higher number of intrusions for poor comprehenders in Experiments 2, 3, and 4 seems to be due to poor suppression mechanisms, rather than to a particular task artifact. In our view, the exclusion of possible intrusions occurs during both the processing of incoming information and the retrieval process involved in giving a response. During presentation, the participant processes each incoming item, attributes appropriate activation to the item, and updates its activation. If an item is wrongly activated or its activation is not updated, it will disturb the process of retrieval of pertinent information and, at worst, will even be remembered erroneously. In fact, an insufficient suppression mechanism makes the number of possible candidates greater, producing a greater probability of remembering nontarget elements.

In conclusion, our data show consistently that the ability to update information in working memory is strictly related to reading comprehension ability. This connection appears to be due to a common underlying control mechanism, the characteristics of which should be investigated further in future research.

REFERENCES

- ANASTASI, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- BADDELEY, A. D. (1990). *Human memory: Theory and practice*. Hove, U.K.: Erlbaum.
- BARNES, J. M., & UNDERWOOD, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology*, **58**, 97-105.
- BROWN, A. L., ARMBRUSTER, B. B., & BAKER, L. (1986). The role of metacognition in reading and studying. In J. Orasanu (Ed.), *Reading comprehension: From research to practice* (pp. 49-75). Hillsdale, NJ: Erlbaum.
- CLARK, H. H., & SENGUL, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition*, **7**, 35-41.
- CONWAY, A. R. A., TUHOLSKI, S. W., SHISLER, R. J., & ENGLE, R. W. (1999). The effect of memory load on negative priming: An individual differences investigation. *Memory & Cognition*, **27**, 1042-1050.
- CORNOLDI, C. (1974). Imagery values for 310 Italian nouns. *Italian Journal of Psychology*, **1**, 211-225.
- CORNOLDI, C., & OAKHILL, J. (1991). *Reading comprehension difficulties: Processes and interventions*. Hove, U.K.: Erlbaum.
- CORNOLDI, C., RIZZO, A., & PRA BALDI, A. (1991). *Prove avanzate MT di comprensione nella lettura* [Advanced MT reading comprehension tests]. Florence: Organizzazioni Speciali.
- DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behaviour*, **19**, 450-466.
- DANEMAN, M., & CARPENTER, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 561-584.
- DANEMAN, M., & MERIKLE, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, **3**, 422-433.
- DE BENI, R., & PALLADINO, P. (2001). *Decay in working memory performances with aging: Intrusion errors analyses*. Manuscript submitted for publication.
- 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*, **51A**, 305-320.
- DE VEGA, M. (1995). Backward updating of mental models during continuous reading of narratives. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 373-385.
- EHRlich, K., & RAYNER, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning & Verbal Behavior*, **22**, 75-87.
- ENGLE, R. W., CANTOR, J., & CARULLO, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 972-992.
- GARNER, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.
- GERNSBACHER, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
- GERNSBACHER, M. A., VARNER, K. R., & FAUST, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 430-445.
- GLENBERG, A. M., & LANGSTON, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory & Language*, **31**, 129-151.
- MAY, C. P., HASHER, L., & KANE, M. J. (1999). The role of interference in memory span. *Memory & Cognition*, **27**, 759-767.
- MCKOON, G., & RATCLIFF, R. (1992). Inferences during reading. *Psychological Review*, **99**, 440-466.
- MORRIS, N., & JONES, D. M. (1990). Memory updating in working memory: The role of central executive. *British Journal of Psychology*, **81**, 111-121.
- MORROW, D. G., BOWER, G. H., & GREENSPAN, S. L. (1989). Updating situation models during comprehension. *Journal of Memory & Language*, **13**, 441-469.
- ORSINI, A., & LAICARDI, C. (1997). *WAIS-R: Contributo alla taratura italiana* [A contribution to Italian standardization]. Florence: Organizzazioni Speciali.
- PASSOLUNGI, M. C., CORNOLDI, C., & DE LIBERTO, S. (1999). Working memory and intrusions of irrelevant information in a group of specific poor problem solvers. *Memory & Cognition*, **27**, 779-790.
- POLLACK, I., JOHNSON, L., & KNAFT, P. (1959). Running memory span. *Journal of Experimental Psychology*, **57**, 137-146.
- ROBERTSON, R. R. W., & GERNSBACHER, M. A. (1996, May). *Suppression during narrative comprehension*. Paper presented at the 68th Annual Meeting of the Midwestern Psychological Association, Chicago.
- VAN DER LINDEN, M., BREDART, S., & BEERTEN, A. (1994). Age-related differences in updating working memory. *British Journal of Psychology*, **85**, 145-152.
- VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- WECHSLER, D. (1987). *Manual for Wechsler Memory Scale-Revised*. New York: Psychological Corporation.