Child Development, July/August 2007, Volume 78, Number 4, Pages 1171 – 1185

Analyzing False Memories in Children With Associative Lists Specific for Their Age

Paula Carneiro
Universidade Lusófona, Portugal

Pedro Albuquerque
Universidade do Minho, Portugal

Angel Fernandez
Universidad de Salamanca, Spain

Francisco Esteves
Universidade Lusófona, Portugal

Two experiments attempted to resolve previous contradictory findings concerning developmental trends in false memories within the Deese-Roediger-McDermott (DRM) paradigm by using an improved methodology– constructing age-appropriate associative lists. The research also extended the DRM paradigm to preschoolers. Experiment 1 (N = 320) included children in three age groups (preschoolers of 3 – 4 years, second-graders of 7 – 8 years, and preadolescents of 11 – 12 years) and adults, and Experiment 2 (N = 64) examined preschoolers and preadolescents. Age-appropriate lists increased false recall. Although preschoolers had fewer false memories than the other age groups, they showed considerable levels of false recall when tested with age-appropriate materials. Results were discussed in terms of fuzzy-trace, source-monitoring, and activation frameworks.

Episodic memories can often be distorted versions of the originally experienced events, with distortions ranging from extensive fabrications to false memories for specific details. Understanding how errors of this kind are produced is a goal shared by many memory researchers, and several empirical approaches are currently used to study different manifestations of memory inaccuracy. In this regard, a particularly productive procedure is exemplified by the Deese-Roediger-McDermott (DRM) paradigm, which has been widely used to study false memories in adults in a way that allows strict levels of experimental control. Originally developed by Deese (1959), and subsequently adapted and extended by Roediger and McDermott (1995), this paradigm provides a straightforward way to study false recall and false recognition through associative processes. It essentially involves presenting participants with lists of related words (e.g., *bed*, *rest*, *awake*, *tired*, *dream*, *wake*, *snooze*, *blanket*, *doze*, *slumber*, *snore*, *nap*, *peace*, *yawn*, *drowsy*) that are, all of them, associates of a nonpresented word–the critical word–(e.g., *sleep*) and subsequently testing them in standard free-recall and recognition tasks. Roediger and McDermott (1995) found that participants recalled the nonpresented critical words at the same (Experiment 1) or higher level (Experiment 2) than the words actually presented in the middle of the lists. In addition, in recognition tasks, the nonpresented critical words were mistakenly identified as being heard approximately as often as words that were actually studied in the lists.

Many studies have replicated these robust false memory effects (e.g., McDermott, 1996; Payne, Elie, Blackwell, & Neuschatz, 1996; Robinson & Roediger, 1997; Stadler, Roediger, & McDermott, 1999), providing evidence that the DRM procedure is a powerful tool for analyzing the production of false memories with a rich variety of manipulations and under rigorous experimental conditions.

Although this paradigm has inspired a substantial number of studies on false memory in different populations, as for instance, older adults (e.g., Balota et al., 1999; Kensinger & Schacter, 1999; Norman & Schacter, 1997) and amnesic patients (e.g., Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, Anes, & Racine, 1998; Schacter, Verfaellie, & Pradere, 1996), its application to children has been infrequent and very recent. Two studies that used the DRM paradigm to analyze the developmental pattern of false memories were published at almost the same time (Brainerd,

This research was supported by a Ph.D. Grant (SFRH/BD/6937/2001) awarded to the first author by the Fundação para a Ciência e a Tecnologia, in Portugal. Additional support was provided to the third author by Junta de Castilla y León (Project SA031A-06) and by the Spanish Ministry of Education and Science (Project SEJ2005-03210). The authors thank Luísa Araújo for her help in translating the lists.

Correspondence concerning this article should be addressed to Maria Paula Carneiro, Universidade Lusófona de Humanidades e Tecnologias, Departamento de Psicologia, Av. do Campo Grande, 376, 1749-024, Lisboa, Portugal. Electronic mail may be sent to mpcarneiro@usal.es.

Reyna, & Forrest, 2002; Ghetti, Qin, & Goodman, 2002), but they showed contrasting results.

The study by Brainerd et al. (2002) reported three experiments in which children of 5, 7, and 11 years of age and adults were tested in a free-recall task. In the last experiment of this study, recall tests were followed by a final recognition task. The results of these three experiments clearly indicated that kindergarten children (age 5) showed a very low proportion of false recall and that there is a developmental pattern, in the sense that false recall and false recognition systematically increased with age (from age 5 to adulthood). According to the authors of the study, these results support the fuzzy-trace theory, which assumes that memory representations involve two separate and distinct traces: a gist trace, which corresponds to the semantic, relational, and elaborative properties of the stimulus and is usually responsible for the false memories, and a verbatim trace, which corresponds to the exact surface form of the stimulus and is responsible for correct recall or recognition and for the correct rejection of distractors. The use of both gist and verbatim information tends to increase with age, with older children being more likely to process gist traces. According to this approach, when conditions favor the processing of gist traces, older children produce more false memories than younger children (Brainerd, Reyna, & Poole, 2000). Since the DRM paradigm involves the presentation of associates that all converge to a critical nonpresented word, gist extraction would be facilitated and becomes a crucial element for the production of false memories. Therefore, and in line with the above arguments, Brainerd et al. (2002) attributed the low level of false memories shown by kindergarten children to their limited gist-processing abilities and the increment of false memories with age to a progressively better ability to process gist as children get older. The results of some more recent studies (Dewhurst & Robinson, 2004; Howe, 2005; Howe, Cicchetti, Toth, & Cerrito, 2004) can also be considered consistent with this view. The studies reported by Howe et al. (2004) and by Howe (2005) replicated this developmental pattern, the first one for both maltreated and nonmaltreated children and the second one with a single DRM list. In addition, Dewhurst and Robinson (2004) showed that semantically related intrusions in recall increase with age, from 5 to 11 years.

However, the study of Ghetti et al. (2002) did not corroborate these results. In their experiment, children of 5 and 7 years of age and adults were tested in both recall and recognition tasks, using the DRM procedure. In this study, the level of false recall was very high even for 5-year-olds, and the level of false memory, both in terms of recall and recognition, remained essentially the same from age 5 through adulthood. The authors of this study also used a measure of relative false recall, which provides an indication of the level of false memory in relation to the total number of recalled words, and for that measure, they found that kindergarten children showed a higher proportion of relative false recall than 7-year-olds and adults. The developmental increase in relative false recall obtained in this experiment was explained by Ghetti et al. (2002) within the source-monitoring framework (Johnson, Hastroudi, & Lindsay, 1993; Johnson & Raye, 1981). According to this theoretical view, false memories arise from the incapacity to correctly attribute the source of the activated information at the time of the test. In the specific case of the DRM paradigm, a false memory would occur when information internally generated, for example, by automatic activation in the semantic network, is wrongly attributed to an external source as, for instance, being previously heard or viewed. Earlier research on children's source-monitoring abilities (Foley & Johnson, 1985; Foley, Johnson, & Raye, 1983) indicated that young children are more likely than adults to confuse memories from different sources when they are self-generated (internal source monitoring). However, Lindsay, Johnson, and Kwon (1991) argued that young children have special difficulties in coping with similarity among sources, regardless of the internal/external distinction. These authors showed that these children also have internal-external reality monitoring difficulties when the sources are highly similar. This could be the case in the DRM paradigm, in which the external source presented by an audiotape does not provide extensive memory information about the actor involved in the event. If this is the case, then young children would have more difficulties in differentiating the internal source from the rather poor external source in the DRM paradigm, and hence they would be more prone to produce false memories.

Solving the disparity between the findings of these two studies regarding the developmental trends in false memories is complicated by the fact that, in addition to their dissimilar theoretical assumptions, they present some potentially important methodological differences that make the comparison difficult. First, the presentation rate differed between the studies. Brainerd et al. (2002) presented words at a rate of 2 seconds, whereas Ghetti et al. (2002) used a 5-second presentation rate. As reported by Gallo and Roediger (2002) slower presentation rates, as occurred in the Ghetti et al. (2002) study, tend to decrease false recall. Thus, this difference could result in higher false recall levels for the Brainerd et al. (2002) study. Second, a retention interval between list presentation and recall was only used by Ghetti et al. (2002). Brainerd et al. (2002) had study lists immediately followed by free recall. According to McDermott (1996), increasing the retention interval could enhance false memory. This would predict more false memories in the Ghetti et al. (2002) study. Third, the number of associates per list was also different. In Brainerd et al. (2002), each to-be-remembered list was composed of 12 associates (Experiment 1) or 15 associates (Experiments 2 and 3), whereas in Ghetti et al. (2002) only 7 associates were presented in each list. Robinson and Roediger (1997) showed more false memories for longer lists than for shorter lists, which could predict more false memories in Brainerd et al. (2002) study. The effects of two of these factors–list length and presentation rate–would predict more false memories for Brainerd et al.'s (2002) study than for Ghetti et al.'s (2002), which in fact did not occur at least for young children. However, as the effects of these three factors on false memories have been identified only for adults, but remain unknown for children, they cannot provide an explanation for the divergent developmental pattern of false memories obtained in the two studies.

Due to the distinct results of these two studies, the present study used a methodology similar to the standard DRM paradigm and focused on the idea that the associative relation between the list words and the critical items was an essential part of the puzzle posed by the empirical findings described above. Thus, as it is regularly done for adults (e.g., Roediger & McDermott, 1995), in the present study the selection of materials was dictated by specific free-association norms appropriate for the participants. The studies that have employed the DRM paradigm with children (Brainerd et al., 2002; Dewhurst & Robinson, 2004; Ghetti et al., 2002; Howe, 2005; Howe et al., 2004) used study lists that were built according to free-association norms originally collected with adult samples. A potential problem with this selection procedure is that young children could be unable to understand the meaning of some of the words. More importantly, and even if only words present in the children's vocabulary are used (as in Ghetti et al., 2002), the extent to which a particular critical word is associated to its corresponding list words is unknown. In order to overcome this problem, children's false memory should be tested with lists that are constructed using free-association norms that are specific for the age of the children. In fact, the studies about associative norms in children have revealed that there can be substantial changes in the type of words and in the hierarchical structure of word associations along the course of development (e.g., Nelson, 1977). In line with this argument, the present study aimed to solve the distinct results produced by previous research, without neglecting the fact that the associative relation in semantic representation is age-specific (Bjorklund, 1987; Nelson, 1977). Two experiments are reported below in which children of different ages were tested with lists that, although sharing the same critical words, were composed of different sets of to-be-remembered associates, as dictated by the specific free-association norms available for each age group. The goal of Experiment 1 was to analyze whether there is a developmental increase in false memories with material appropriate for each age. Moreover, it studied the DRM paradigm in preschoolers for the first time. False recall and false recognition were assessed in three different groups of children (preschoolers, second-graders, and preadolescents) and one group of adults, with lists of associates adapted to age. In Experiment 2 the developmental pattern of false memories was further studied by using selected lists that produced more false memories for children. Furthermore, the effect of age-specific associative lists was also analyzed with age-matched and age-mismatched lists of associates, in order to examine whether this was a crucial variable that increased false memories in children.

Experiment 1

The purpose of the first experiment was to study the developmental pattern of false memories using the DRM paradigm with associative lists specific to the participant's age. Although the two main studies that analyzed false memories in children (Brainerd et al., 2002; Ghetti et al., 2002) showed discrepant results, the other three more recent studies (Dewhurst & Robinson, 2004; Howe, 2005; Howe et al., 2004) corroborate the Brainerd et al. (2002) findings and seem to predict an increase of false memories with age. However, all these studies used adult lists to test children, which could have produced a decrease in the false memory levels of preschoolers, because the presented lists did not promote gist traces for them. The present study would predict that appropriate material would increase their false memories, and differences between children and adults could be reduced substantially relative to previous studies.

With respect to the methodological aspects of rate of presentation and retention interval, this experiment followed the same procedures as in Roediger and McDermott (1995). In these aspects, this experiment is more similar to the Brainerd et al. (2002) study than to the Ghetti et al. (2002) study. However, in relation to list length, other findings about memory development were taken into consideration to establish the number of associates per list in each age group. List length was adjusted between groups according to the development of memory span (Dempster, 1981) and to the evidence of substantial age differences in the performance of free-recall tasks (Kail, 1990; Schneider & Bjorklund, 1998).

Method

Participants. A total of 320 participants, 256 children and 64 young adults, took part in this experiment. Children belonged to three age groups: 128 preschoolers (mean age = 4 years, 2 months; range = 3y, 0m to 4y, 11m; 64 female and 64 male participants), 64 second-graders (mean age = 7 years, 9 months; range = 7y, 0m to 8y, 10m; 26 female and 38 male participants), and 64 preadolescents (mean age = 12 years, 5 months; range = 11y, 10m to 12y, 10m; 32 female and 32 male participants). The young adults were undergraduate volunteers (45 female and 19 male participants) with ages ranging from 18 to 34 years and a mean age of 24 years. All the participants were native speakers of Portuguese, mostly from middle-class backgrounds, and they were recruited in the nurseries, elementary schools, and universities of Lisbon (Portugal).

Materials. Sixteen critical words and their corresponding age-specific study lists were selected on the basis of normative data for words in Portuguese (Carneiro, Albuquerque, Fernandez, & Esteves, 2004). The critical words, common for the four age groups, were frequent concrete nouns already present in the vocabulary of preschool children. A separate set of 16 study lists was constructed for each age group (see the lists' translation in Appendix A. The Portuguese words and the association levels of the critical words are available from the authors). For any given group, the study list corresponding to a particular critical word was formed by the words most frequently associated to it, with items in each list placed in decreasing order of associative frequency, according to previously collected free-association norms that were specific to the age of the group members (Carneiro et al., 2004). List length varied between groups: 8, 10, 12, and 15 items for preschoolers, second graders, preadolescents, and adults, respectively. Thus, lists for all groups shared the same critical words, but they differed quantitatively in terms of number of items and qualitatively in terms of some of the particular words included and their serial position. Although the lists selected for children were shorter than the lists selected for adults, they were structurally equivalent and very similar regarding the total associative strength of the connections between critical words and the items in their corresponding lists (an average of 0.71 in the children's lists, and of 0.72 in the adults' lists). In order to keep the experimental sessions short for the preschoolers, each child was tested on four lists, compared to eight lists in the remaining groups. A counterbalance procedure was used to make sure that all the lists in a given age-specific set were used an equivalent number of times (32).

The recognition test included 24 words for the preschoolers and 48 words for the remaining groups. For each participant, half the words in the test were selected from the presented lists (two studied words, from the initial and the middle serial positions, and the critical word from each list) and the other half were unrelated distractors consisting of initial items, middle items, and critical words from the lists that were not presented. The order of the words in the recognition test was randomly determined for each participant.

Procedure. All the participants were tested individually. They were told that they would hear some words played on a tape recorder and then they would be asked to recall all the words they heard, in any order. Participants were advised to recall only the words they heard but as many words as they could. Instructions were adapted to the participants' age, and for the youngest group, the task took the form of a memory game. The lists were previously recorded on tape, and words were presented at a rate of one item every 2 seconds. After listening to each list, participants had approximately 2 minutes to say the words they remembered while the experimenter wrote them down on a sheet of paper and recorded the participants' oral responses on an audiocassette, in order to allow a later recovery of any response that could be missed on-line. The listening and oral recall sequence was repeated until the last list was presented. The presentation order of the lists was counterbalanced across the participants within each group.

After recalling the items from the last list, participants were administered a final recognition test. They were instructed to listen to the words spoken by the experimenter and, for each word, say whether they remembered hearing the word earlier, on the tape recorder, or not. Responses were registered by the experimenter on a score sheet. Instructions for preschoolers emphasized that some of the words were actually heard, whereas others were not. Even so, eight preschool children gave the same response to all the words in their test, and their recognition data (just over 6% of the total) were excluded from the analyses.

Results

Recall and recognition data are presented in Table 1 (see also the percentages of correct and false recall for each list at each age in Appendix B). Regarding the recall data, the proportions of studied words (correct recall) were calculated dividing the total number of correctly recalled words by the number of studied lists and the number of items presented in each list, and the proportions of critical words (false recall) were obtained dividing the number of critical words falsely recalled by the number of studied lists. The proportion of intrusions (all the other words recalled that were neither studied nor critical) was calculated in relation to the number of studied lists and the total words recalled for each list.

A preliminary analysis of gender indicated no effect on recall or recognition scores. Thus, gender was not included in further analyses.

Free Recall

Correct recall. The results of a one-way ANOVA revealed a significant effect of age group on the proportion of correct recall, $F(3,316) = 194.70$, $MSE = .01$, $p < .001$. Post-hoc analyses (Tukey tests in this and subsequent analyses) showed that, in children, correct recall reliably increased with chronological age, with significant differences found among all the children groups (all ps $< .001$). However, the comparison between preadolescents and adults did not show a significant difference.

Figure 1 presents the serial position curves for correctly recalled words in each age group. In order to get smoothed curves, the data points for each position were calculated by averaging the mean value of a position plus the mean values of the two adjacent positions. For the first and the last positions, the values correspond only to the means of those positions. As can be seen in the figure, a standard recency effect is observed in all age groups, whereas a primacy effect is almost nonexistent in the group of preschoolers. The relevance of this result for the understanding of false memories formation will be addressed in the Discussion section.

False recall. A 2 $\times$ 2 ANOVA examined age group (preschoolers, second-graders, preadolescents, and adults) and type of item (criticals vs. intrusions). A significant main effect of age group, $F(1,316) = 4.37$, $MSE = .01$, $p < .01$, indicated that preschoolers recalled, in general, lower levels of nonpresented items than any other age group ($p < .05$ in all the comparisons). Also, a main effect of type of item, $F(1,316) = 195.09$, $MSE = .02$, $p < .001$, showed that the proportion of critical items was higher than the proportion of intrusions. However, the significant interaction between age group and type of item, $F(3,316) = 42.03$, $p < .001$, clarified that false recall for critical items increased from preschoolers to all the other age groups, whereas the recall of intrusions decreased. Additional t-tests for each age group revealed that the proportion of critical items recalled was higher than the proportion of intrusions for second-graders, $t(63) = 7.48$, $p < .001$, for preadolescents, $t(63) = 10.84$, $p < .001$, and for adults, $t(63) = 8.49$, $p = .001$. In the case of preschoolers, no significant differences in the recall of these two different types of items were found. Thus, the preschoolers showed the same level of false recall for both types of items, whereas the other groups' age showed considerably more false recall for the critical items.

In order to analyze whether the intrusions were related in meaning to the other presented items,

Table 1
Proportions of Different Types of Responses on the Recall and Recognition Tasks of Experiment 1

			Preschoolers	Second-graders	Preadolescents	Adults
Recall	Studied		.31	.49	.62	.60
	Critical		.09	.21	.25	.24
	Intrusions	T	.11	.06	.02	.03
		R	.05	.04	.02	.03
		NR	.06	.02	.00	.00
Recognition	Studied		.70 (.82)	.71 (.90)	.84 (.94)	.88 (.95)
	Critical		.52 (.72)	.57 (.84)	.56 (.83)	.60 (.85)
	Distractors		.16	.03	.02	.03

Notes. Data in parentheses are A′ values.
T-totals; R-related; NR-not related.

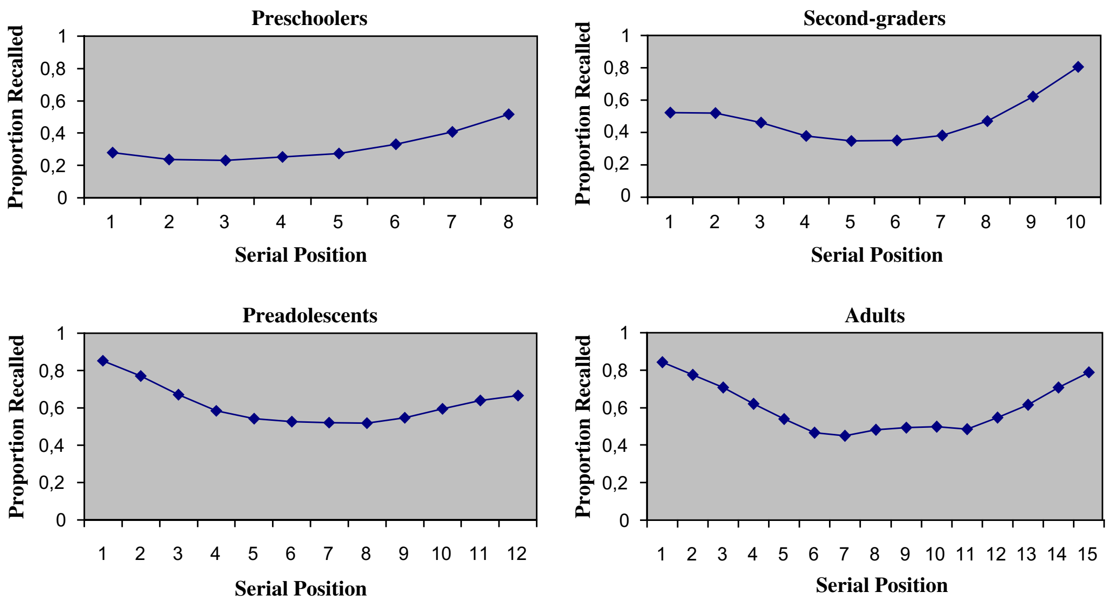

Figure 1. Serial position of correctly recalled items for the different age groups in Experiment 1.

which could suggest another type of false recall not so restricted to the critical items, further analyses were conducted. The intrusions were classified into related and not related, by two independent examiners, with a coefficient of agreement above 90%. Related intrusions were words not presented in a given list but similar in meaning to that list's studied words, excluding the critical item, whereas unrelated intrusions were recalled words that had no meaning resemblance to the list's studied items, including items that belonged to previously presented lists. The comparative analyses of age group for these two types of intrusions showed that preschoolers recalled significantly more unrelated intrusions than the other age groups, $F(3,316) = 18.70$, $MSE = .004$, $p < .001$, and also more related intrusions than preadolescents, $F(3,316) = 4.78$, $MSE = .003$, $p = .01$, (see Table 1). Thus, preschoolers showed more intrusions than all the other age groups, regardless of whether they were related or not related to the studied words.

As already described, preschoolers had the lowest recall levels for critical words. However, this could be due to the fact that overall they tended to produce fewer items in the recall task. In order to analyze false recall in relation to the total of words recalled (a relative false recall measure), for each participant, the number of recalled critical lures divided by the total number of recalled words produced a proportion score. (This proportion score differs from the one above, which we based on total words presented rather than total words recalled.) Mean relative false recall was .03 for preschoolers, .04 for second graders, .03 for preadolescents, and .03 for adults. A one-way ANOVA showed no significant effect of age group for relative false recall.

Usually, the measure of false memories in the adult studies about the DRM paradigm is restricted to the critical items. Nevertheless, for completeness, an analysis on a broader measure of false recall was also conducted, adding intrusions to critical lures and dividing by the proportion of total words recalled. This combined measure provides an indication of the production of unstudied words of any kind, taking into account the specific recall capacity of each age group. Overall, this index of incorrect recall showed a developmental pattern, and it was higher for preschoolers (.15) than for the rest of the groups (.10, .05, and .06 for second-graders, preadolescents, and adults, respectively). A one-way ANOVA revealed a significant effect of age group, $F(3,316) = 17.89$, $MSE = .01$, $p < .001$, and the post-hoc tests found significant differences between the preschoolers and all the other age groups ($p < .05$ in comparison to second-graders and $p < .001$ in comparison to the other age groups) and between second-graders and preadolescents or adults ($p < .05$). The same kind of analysis performed with only the related intrusions showed a significant age difference between the preschoolers (.05) and the adults (.03), $F(3,316) = 3.03$, $MSE = .002$, $p < .05$.

In short, preschool children showed the lowest levels of false recall but no difference when the total words produced were taken into account. Also, a broader measure of false recall, including all type of intrusions or only the related intrusions, revealed a higher level for preschoolers.

Recognition. The mean proportions of "yes" responses to studied, critical, and distractor words are presented in Table 1. Preliminary analyses showed that the rate of false alarms was not similar across age groups, $F(3,308) = 25.11$, $MSE = .02$, $p < .001$, and pair-wise comparisons further indicated that the rate of false alarms in preschoolers was significantly higher than that of all other age groups (all $ps < .01$), which replicates previous studies (see Fritzley & Lee, 2003). Because of this discrepancy in false alarms, recognition data were analyzed by signal detection theory (SDT). Therefore, true recognition and false recognition scores were corrected for differences in false alarms through A', a nonparametric counterpart of the signal detection statistic d' (Macmillan & Creelman, 2005). Before applying the SDT formulas, the values of studied or hits (H), nonstudied distractors or false alarms (FAH), criticals (C), and critical distractors of unstudied lists or false alarms to criticals (FAC) were transformed to avoid values of 0 or 1, by adding .5 to each frequency and dividing by N + 1. To obtain the corrected target measure, if H > FAH, the nonparametric formula (A') applied was: $A' = .5 + [(H - FAH)(1 + H - FAH)] / [4 H (1 - FAH)]$, but if H< FAH, an alternative formula was applied: $A' = .5 - [(FAH - H)(1 - H + FAH)] / [4 FAH (1 - H)$. To obtain the corrected false measure, if C > FAC, the formula applied was: $A' = .5 + [(C - FAC)(1 + C - FAC)] / [4 C (1 - FAC)]$, but if C< FAC, then $A' = .5 - [(FAC - C)(1 - C + FAC)] / [4 FAC (1 - C)$. The A' values for studied and critical items are presented in Table 1.

Two separate one-way ANOVAs were performed for correct recognition and false recognition. With respect to true recognition, a significant main effect of age was found, $F(3,308) = 46.18$, $MSE = .007$, $p < .001$. Planned comparisons showed significant higher levels of correct recognition with increasing age ($ps < .05$), except for the comparison between preadolescents and adults. Regarding false recognition, a significant main effect of age also occurred, $F(3,308) = 22.39$, $MSE = .02$, $p < .001$, with preschoolers showing the lowest levels in comparison to all the other age groups ($ps < .001$).

Discussion

Previous studies have analyzed false memory in children with the DRM paradigm using associative lists based on free association norms for adults, producing inconsistent results. The present study, in which younger participants were included and age-specific lists were used, has produced clearer evidence of a developmental trend in false memories, contributing to resolve previous conflicting findings.

The results of this experiment, the first to test preschoolers on the DRM paradigm, revealed significant differences in false recall and in false recognition, as a function of age, mainly between preschool children and the other age groups. Preschoolers showed the lowest levels of false recall, and they also demonstrated the lowest levels of false recognition. Because children were tested with lists composed of associates derived from age-specific norms, these results demonstrate that preschool children are in fact less prone to the DRM memory illusion than older children or adults, a result reported by Brainerd et al. (2002).

As already mentioned, the resistance of preschoolers to this type of illusion has been explained by fuzzy-trace theory (Brainerd et al., 2002). According to this view, the gist representation of the studied verbal material is the factor responsible for the creation of false memories in the DRM paradigm, and efficient processing of the gist trace systematically improves with age. Accordingly, preschoolers should show fewer false memories with DRM paradigm than older children or adults. Additional supporting evidence comes from studies that show that spontaneous use of semantic organization strategies (e.g., elaboration) in recall tasks is common in adolescents and adults but not in preschoolers (Bjorklund & Douglas, 1997; Kail, 1990). Bjorklund and Hock (1982) have argued that younger children processed the words presented in a list on an item-by-item basis, without establishing relationships between the items. Furthermore, the serial position curves of the free-recall task in the present study showed a primacy effect for all age groups except preschoolers. Others who have reported such results (e.g., Bjorklund & Muir, 1988) have argued that the primacy effect occurs only later in development because of increased use of memory strategies. Thus, preschoolers that have difficulties in performing some kind of elaborative rehearsal could prevent the establishment of interitem associations needed for the formation of gist representations of the lists.

But do these results mean that preschoolers are unable to process gist? Preschoolers' lower level of false memories may simply reflect an inability to process gist or just a difficulty that could be overcome under some circumstances. We had hypothesized that gist failure could be reduced by using appropriate material. Our preschoolers (ages 3–4) generated considerable false recall (.09), which was greater than that of the 5-year-old kindergarteners (.06 and .05) in two experiments by Brainerd et al. (2002). This difference probably reflects our use of age-based association norms. Such materials could facilitate gist-processing operations, and even young children could show higher levels of false memories.

An analysis of the variability of false recall with the different associative lists also suggests that gist processing can be facilitated under some circumstances. It is well known that, for adult samples, some lists systematically produce low levels of false recall, whereas others produce high levels (e.g., *king* vs. *window* in Stadler, Roediger, & McDermott, 1999). In the present experiment, preschool children also showed a great variability of false recall. Their proportion of false recall for the 16 lists ranged from 0 to .31. Additionally, in one of the studied lists (the one for the critical word *dog*) the proportion of false recall exceeded the proportion of true recall (.22 vs. .20). Thus, although young children generally showed less false memory than older children, for some of the lists, they could produce quite a lot of false recall.

From a different point of view, the results for false recall also suggest that using appropriate material can substantially reduce age differences. When the number of total words recalled is taken into consideration, there are no age differences in the level of relative false recall. It is possible that age-related differences in false recall of critical lures simply reflect more general age-related differences in the number of items they produce in the recall task rather than an inability to process gist or to process interitem relations. Although previous studies did not provide the values for this measure (except Ghetti et al., 2002), this measure should be used in future research, because it could enrich current interpretations of false recall results.

The smaller age differences in the present study, compared to the results of Brainerd et al. (2002), require some comment. In the present study, no significant increment in false recall or false recognition was found between second-graders and preadolescents, and preadolescents did not differ from adults in correct or false recall/recognition. However, these findings should be interpreted with caution, because, despite the fact that true recall was analyzed in proportion to the number of items in each list, the number of items per list differed in the age groups. It could be argued that the different number of associates per list could affect the tendency to form memory illusions to the critical items. This issue was addressed in the following experiment.

Experiment 2

This experiment focused on two aspects of the lists used in Experiment 1—age-specific associations and list length. First, did the use of age-specific lists actually produce higher levels of false memories? In Experiment 2, lists that were specific to preschoolers and lists that were specific to preadolescents were used with both age groups. If the administration of age-specific materials determines the production of false recall and recognition in the DRM paradigm, participants should show higher levels of false memory when the materials are suited to their age group than a different age group.

Second, because younger children have a smaller memory span (Dempster, 1981) and show lower free recall (Kail, 1990; Schneider & Bjorklund, 1998), in Experiment 1 it seemed reasonable to adapt the number of words presented to the specific memory abilities of the children. However, this length-adjustment procedure could raise another kind of problem. There is considerable evidence from studies with adults that the more associates presented in a study list, the more false recall and false recognition occur (Hall & Kozloff, 1973; Hintzman, 1988; Robinson & Roediger, 1997). If children and adults are equally prone to the effect of list length in the production of false memories, then it could be argued that preschoolers' relative low rates of false recall and false recognition could be due to their shorter lists. Thus, Experiment 2 used shorter lists (with eight associates) and longer lists (with 12 associates) in both preschoolers and preadolescents. Lengths of eight and 12 associates were chosen, because they correspond to the list length of preschoolers and preadolescents, respectively, in Experiment 1.

This experiment included the lists that produced higher levels of false recall for preschoolers and preadolescents in Experiment 1. These lists should increase false memories, confirming that young children can process semantic gist when appropriate materials are employed.

Method

Participants. Participants included 32 preschool children aged 3–4 (mean age = 4 years, 5 months; range = 3y, 4m to 4y, 11m; 16 female and 16 male participants) and 32 preadolescents aged 11-12 (mean age = 11 years, 7 months; range = 11y, 3m to 12y, 11m; 16 female and 16 male participants). All were Portuguese native speakers and were recruited from an upper-class private school. None took part in Experiment 1.

Materials and procedure. Four lists that produced the highest levels of false recall for both preschoolers and preadolescents were selected from Experiment 1 (see Appendix A–Lists of Experiment 2). For each participant, four lists were prepared: one short list specific for preschoolers, one short list specific for preadolescents, one long list specific for preschoolers, and one long list specific for preadolescents. Short lists included eight associates, and long lists comprised 12 associates. The different lists were counterbalanced between these four conditions to ensure that each list was presented in each condition an equal number of times.

For all participants, the recognition test was composed of 24 words: 12 studied words from the four presented lists (items in positions 1, 3, and 5 in each list); and 12 words that were not studied, with four of them being the critical items of the studied lists and the other eight being distractors (three associates and one critical item from each of two unpresented lists). The same procedure as in Experiment 1 was used, except that each participant received only four lists.

Results

The memory measures for studied words, critical words, intrusions, and distractors were calculated in the same way as in Experiment 1, and the corresponding means for each experimental condition are presented in Table 2. Independent mixed ANOVAs analyzed the effects of list type (specific to preschoolers vs. specific to adolescents), list length (8 or 12 items), and age group (preschoolers and preadolescents) on each recall measure.

Free Recall

Correct recall and intrusions. As in Experiment 1, and in line with standard findings in this area, preadolescents recalled a significant higher proportion of studied words (M = .62) than preschoolers (M = .43), $F(1,62) = 64.45$, $MSE = .04$, $p < .001$, and shorter lists produced significantly better correct recall (M = .59) than longer lists (M = .46), $F(1,62) = 128.28$, $MSE = .009$, $p < .001$. Regarding intrusions other than the critical words, their number was very low, and the only noticeable pattern was a near-significant tendency for longer lists to produce more intrusions when they were of the preschooler type than when they were of the adolescent type (for this interaction between list length and list type, $F(1,62) = 3.32$, $MSE = .004$, $p = .07$).

False recall. Preliminary analyses showed that the proportion of critical words recalled was considerably higher than the proportion of unrelated intrusions–.26 and .03 for preschoolers, $t(31) = 5.64$, $p < .001$; and .41 and .02 for preadolescents, $t(31) = 7.56$, $p < .001$. The analyses that focused on the specific effects of the variables of interest on the recall of the critical items showed a main effect of age group, $F(1,62) = 5.43$, $MSE = .26$, $p < .05$, with preadolescents recalling a significantly higher proportion of critical words recalled (M = .41) than the preschoolers (M = .26). In addition, and importantly, an interaction effect between age group and list type, $F(1,62) = 5.56$, $MSE = .20$, $p < .05$, reflected the fact that 3-4-year-olds produced more false recall when the lists were specific to preschoolers, whereas 11-12-year-olds recalled more critical words when the lists were specific to preadolescents.

As in Experiment 1, a relative false recall measure was obtained for each age group in which the proportion of false recall was divided by the proportion of total words recalled. This 2 × 2 × 2 ANOVA also showed an interaction effect between age group and

Table 2
Proportions of Different Types of Responses on the Recall and Recognition Tasks of Experiment 2

		Preschoolers	Preadolescents
Recall			
Studied			
Preschooler lists	Short	.50	.69
	Long	.34	.55
Preadolescent lists	Short	.51	.65
	Long	.36	.57
Critical			
Preschooler lists	Short	.31	.34
	Long	.31	.31
Preadolescent lists	Short	.19	.47
	Long	.22	.50
Intrusions			
Preschooler lists	Short	.01	.03
	Long	.05	.03
Preadolescent lists	Short	.03	.02
	Long	.01	.02
Recognition			
Studied			
Preschooler lists	Short	.91 (.89)	.91 (.91)
	Long	.81 (.86)	.79 (.88)
Preadolescent lists	Short	.88 (.87)	.88 (.90)
	Long	.83 (.86)	.80 (.89)
Critical			
Preschooler lists	Short	.63 (.74)	.72 (.79)
	Long	.53 (.72)	.66 (.77)
Preadolescent lists	Short	.81 (.80)	.69 (.78)
	Long	.66 (.75)	.75 (.79)

Note. Data in parentheses are A′ values.

list type, $F(1,62) = 5.76$, $MSE = .01$, $p < .05$. This result revealed that, even when the total words recalled were taken into consideration, 3–4 year-olds produced more false recall for lists specific to preschoolers than to preadolescents ($M = .06$ vs. $M = .04$), whereas 11–12 year-olds produced more false recall for lists appropriate for them ($M = .08$) rather than for preschoolers ($M = .05$).

Recognition. A preliminary comparison for independent groups revealed that preschoolers wrongly recognized a significant higher proportion of distractors than preadolescents [$M = .10$ vs. $M = .03$; $t(62) = 2.77$, $p <. 01$]. As in Experiment 1, the differential false-alarm rate called for the application of a score transformation procedure based on signal detection theory, using nonparametric A' values. Mean proportions and A' values for the different conditions are presented in Table 2. Two separate mixed ANOVAs assessed the effects of age group, list type, and list length on correct recognition and on false recognition A' scores. In regard to correct recognition, the results showed a main effect of list length, $F(1,62) = 7.97$, $MSE = .004$, $p < .01$, with shorter lists significantly better recognized ($M = .90$) than longer lists ($M = .87$). Correct recognition values tended to be higher for preadolescents ($M = .90$) than for preschoolers ($M = .87$), but this difference was marginally significant, $F(1,62) = 3.12$, $MSE = .01$, $p = .08$. In the case of false recognition performance, the analyses did not reveal any reliable main or interaction effects.

Discussion

In general, the results corroborated the age differences in recall in Experiment 1. Preadolescents revealed a higher proportion of correct and false recall than preschoolers and, although not significant, a tendency for a lower proportion of intrusions. In Experiment 2, both false recall and false recognition were considerably higher than in Experiment 1, consistent with the fact that the lists selected for Experiment 2 were those that produced more false recall and false recognition in Experiment 1. Probably for the same reason, the proportion of critical words recalled was significantly higher than the proportion of intrusions for both age groups, which meant that, even for preschoolers, a considerable amount of false recall was found. It is interesting to note that these lists that produced higher levels of false recall also produced lower levels of intrusions. Although it is possible that the lower intrusion level could, at least in part, be related to the fact that participants studied fewer lists in Experiment 2, this finding indicated that these optimized lists facilitated the formation of a gist representation particularly consistent with the critical word. Such gist processing may have prevented a less focused type of processing that would produce other intrusions.

The recognition data did not reveal a significant effect of age group for false recognition. Compared to recognition, the recall task was more sensitive to differences of age, which, according to some authors, supports the notion that the performance of younger children is more impaired the more the task requires the use of strategies (e.g., Bjorklund & Coyle, 1995; Cox, Ornstein, Naus, Maxfield, & Zimler, 1989; Schneider & Bjorklund, 1998).

Regarding the main aim of this experiment, the results clearly demonstrated that children produce higher levels of false recall with lists specially built for their own age. This was shown by an interaction effect between age group and list type for both simple and relative false recall measures. This finding is particularly important, because it justifies the procedure used in Experiment 1 and, compared to the previous studies in this topic (Brainerd et al., 2002; Dewhurst & Robinson, 2004; Ghetti et al., 2002), it provides a better way to study false recall in children. Note that the results of the two experiments demonstrate that younger children can show substantial false recall when tested with age-appropriate materials, and, at the same time, they provide evidence of lower levels of false recall in preschoolers in comparison to older children.

The other purpose of this experiment was to study the effect of list length with children. The results showed that shorter lists produced higher levels of correct recall and correct recognition than longer lists, which is consistent with the results obtained by Robinson and Roediger (1997) with adults. However, the present study did not confirm the expectation that for children, longer lists would produce higher levels of false recall than shorter lists. This absence of a list-length effect in the present experiment suggests that the findings of Experiment 1 regarding the developmental increase of false memories are not attributable to the fact that older participants studied longer lists. And, from a broader perspective, this result could mean that for children, at least until preadolescence, the number of associates in a list it is not a crucial factor for false recall. However, this result should be interpreted with caution, since the contrast between 8 and 12 items in list length could have been insufficient to make a difference.

General Discussion

The two experiments clearly indicate that both children and adults are prone to the DRM false-memory effect when associative lists are specially built for their specific ages. However, the results also showed that, in general, preschool children create lower levels of false memories than older children, replicating the developmental pattern of false memories observed by Brainerd et al. (2002). Additionally, the recall task proved to be more sensitive to age differences than the recognition task. For instance, Experiment 2 failed to reveal age differences in false recognition, probably because the lists used were specially chosen to produce false memories in children.

To explain the pattern of false memories in preschoolers, different theoretical frameworks could account for the observed results. According to the fuzzy-trace theory (Brainerd et al., 2002), the gist representation of preschoolers is yet immature, that is, these children have particular difficulties in extracting the gist of the associative lists, and hence they produce fewer false memories.

But an activation-based theory could also provide another kind of explanation for these results. Although not yet explicitly considered by the proponents of activation approaches (e.g., Roediger, Balota, & Watson, 2001; Underwood, 1965), it seems reasonable to pay attention to the development of the semantic network and to the evolution of semantic memory relations in order to account for the observed age differences in false memory. According to Bjorklund (1987), if knowledge is represented as a semantic network composed of connecting nodes, it is likely that, with development, the number of nodes increases, as does the strength of the connections between them, and the ease of their activation. With an enlarged knowledge base, and broader experiences, the semantic memory representations of items may have more features integrated into them, more readily activated connections with other items, and lower thresholds for their activation. These changes pave the way for the use of more efficient retrieval cues. According to this view, the higher the likelihood that interitem relations are activated, the better the performance in memory tasks. In this argument, the lower spreading activation levels in preschoolers could result in less activation of the critical words, and, consequently, lower rates of false memory in the DRM paradigm. In other words, age differences in false recall could be the result of age-related differences in the nature of activations.

An alternative theoretical view is one based on the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981). As described in the introduction, when two memory sources cannot be easily discriminated, as in the DRM paradigm (hearing some words vs. thinking of them) children's ability to correctly attribute a source to experienced events might increase with age. This leads to the general prediction that younger children would be likely to confuse false memories (thoughts) with real memories (studied words) and, as a result, would produce more false memories to critical items than older children. But the two experiments reported here showed less false memory in young children, suggesting that inefficient use of monitoring strategies cannot, on its own, explain the developmental pattern of false memories. In order to adequately account for these results, the operation of source-monitoring mechanisms must be related to the initial encoding conditions, as suggested by Johnson et al. (1993), and as more concretely exemplified by the activation/monitoring theory (Roediger, Balota, & Watson, 2001). This recent integrative approach would predict that once a nonpresented item is activated, younger children would find it harder to edit out or suppress the false memory, because they cannot rely on source information to distinguish true from false memories. Actually, the fact that preschoolers recalled a large number of related and unrelated intrusions in Experiment 1 is highly consistent with this line of argumentation, and it is likely that further research along these lines would be fruitful.

Independently of the particular theoretical assumptions, it is plausible that at least two kinds of processes could operate in false memory development with the DRM paradigm. On one hand, the application of strategic processes, such as elaboration and rehearsal, could facilitate the extraction of gist. On the other hand, an enriched knowledge base could increase spreading activation to critical words, without the need of conscious identification of the gist. Both processes seem to develop with age (Schneider & Bjorklund, 1998).

One of the most striking findings of this study is that using associative lists specific to children's age increases false recall. Experiment 2 confirmed that the use of age-specific lists in Experiment 1 is more appropriate than the use of adult lists, the practice in previous studies (Brainerd et al., 2002; Ghetti et al., 2002). The present study highlights the associative nature of the lists as one of the great determinants of children's false memory. The finding that, making use of age-specific free-association norms, children can show higher levels of false memories means that children are not so poor in gist processing abilities as previously assumed. Although young children seem to process gist to a lesser extent than older children or adults, they are able to process it if appropriate material is provided. An interesting focus for further studies would be a detailed analysis of the characteristics of the specific lists that create elevated levels of false memories in an age group that, with unmodified standard verbal materials, tends to produce low levels of false memories.

References

Balota, D. A., et al. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Neuropsychology, 16*, 361 – 384.

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretative review. *Developmental Review, 7*, 93 – 130.

Bjorklund, D. F., & Coyle, T. R. (1995). Utilization deficiencies in the development of memory strategies. In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies* (pp. 161 – 180). Hillsdale, NJ: Erlbaum.

Bjorklund, D. F., & Douglas, R. N. (1997). The development of memory strategies. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 201 – 246). Hove, UK: Psychology Press.

Bjorklund, D. F., & Hock, H. S. (1982). Age differences in the temporal locus of memory organization in children's recall. *Journal of Experimental Child Psychology, 32*, 347 – 362.

Bjorklund, D. F., & Muir, J. E. (1988). Children's development of free recall memory: Remembering on their own. *Annals of Child Development*, *5*, 79 – 123.

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development*, *73*, 1363 – 1377.

Brainerd, C. J., Reyna, V. F., & Poole, D. A. (2000). Fuzzy-trace theory and false memory: Memory theory in the courtroom. In D. F. Bjorklund (Ed.), *False-memory creation in children and adults: Theory, research, and implications* (pp. 93 – 127). Mahwah, NJ: Erlbaum.

Carneiro, P., Albuquerque, P., Fernandez, A., & Esteves, F. (2004). Normas de associação livre de 16 palavras portuguesas para crianças de diferentes faixas etárias. *Laboratório de Psicologia*, *2*, 49 – 76.

Cox, B. D., Ornstein, P. A., Naus, M. J., Maxfield, D., & Zimler, J. (1989). Children's concurrent use of rehearsal and organizational strategies. *Developmental Psychology, 25*, 619 – 627.

Deese, J. (1959) On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17 – 22.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, *89*, 63 – 100.

Dewhurst, S. A., & Robinson, C. A. (2004). False memories in children: Evidence for a shift from phonological to semantic associations. *Psychological Science*, *15*, 782 – 786.

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development*, *56*, 1145 – 1155.

Foley, M. A., Johnson, M. K., & Raye C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development*, *54*, 51 – 60.

Fritzley, V. H., & Lee, K. (2003). Do young children always say yes to yes-no questions? A metadevelopmental study of the affirmation bias. *Child Development*, *74*, 1297 – 1313.

Gallo, D. A., & Roediger, H. L. (2002) Variability among word lists in eliciting memory illusions: evidence for associative activation and monitoring. *Journal of Memory and Language*, *47*, 469 – 497.

Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology, 38*, 705 – 718.

Hall, J. W., & Kozloff, E. E. (1973). False recognitions of associates of converging versus repeated words. *American Journal of Psychology, 86*, 133 – 139.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528 – 551.

Howe, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science*, *16*, 927 – 931.

Howe, M. L., Cicchetti, D., Toth, S. L., & Cerrito, B. M. (2004). True and false memories in maltreated children. *Child Development*, *75*, 1402 – 1417.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3 – 28.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67 – 85.

Kail, R. (1990). *The development of memory in children* (3rd ed.). New York: Freeman.

Kensinger, E. A., & Schacter, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology, 16*, 399 – 415.

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52*, 297 – 318.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. (2nd ed.). Mahwah, N.J.: Erlbaum.

McDermott, K. B. (1996) The persistence of false memories in list recall. *Journal of Memory and Language*, *35*, 212 – 230.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin*, *84*, 93 – 116.

Norman, K. A., & Schacter, D. L. (1997) False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, *25*, 838 – 848.

Payne, D. G., Ellie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996) Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language*, *35*, 261 – 285.

Robinson, K. J., & Roediger, H. L. (1997) Associative processes in false recall and recognition. *Psychological Science*, *8*, 231 – 237.

Roediger, H. L., Balota, D. H., & Watson, J. M. (2001). Spreading activation and the arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath & A. M.

Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington, DC: American Psychological Association.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814.

Schacter, D. L., Verfaellie, M., & Anes, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology, 11*, 331–342.

Schacter, D. L., Verfaellie, M., Anes, M. D., & Racine, C. (1998). When true recognition supresses false recognition: Evidence from amnesic patients. *Journal of Cognitive Neuroscience*, *10*, 668–679.

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996) The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language*, *35*, 319–334.

Schneider, W., & Bjorklund, D. F. (1998). Memory. In W. Damon (Series Ed.), D. Kuhn & R.S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 467–521). NewYork: Wiley.

Stadler, M. A., Roediger, H. L., & McDermott, K. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*, 494–500.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122–129.

Appendix A

Lists of Experiment 1

Preschoolers

Airplane: fly, wings, sky, air, fall, travel, people, helicopter

Bed: sleep, pillow, sheets, lay, ni-night, wake up, crib, mother

Book: read, see, stories, paint, drawings, to count, note-book, pages

Car: wheels, drive, steering-wheel, windows, doors, father, glass, lights

Clothing: dressing, skirt, pants, sweater, socks, undress, coat, wash

Dog: ruff ruff, bark, bite, feet, lick, puppy, bones, ears

Door: open, close, key, enter, house, lock, white, wood

Face: eyes, mouth, nose, hair, ears, cheeks, head, beautiful

Mother: good, sleep, play, father, kiss, house, lap, scold

Rain: water, wet, fall, drops, umbrella, sky, clouds, cold

Song: sing, dance, (a synonym of song), radio, music, listen, play, tape

Stone: hard, floor, fall, hurt, beach, (a synonym of hurt), throw, heavy

Street: cars, road, walk, sidewalk, houses, crosswalk, (a synonym of walk), coach

Tooth: eat, mouth, white, chew, wash, broken, dirty, fall

Tree: leaves, fruits, log, apples, flowers, green, garden, street

Water: drink, cold, bath, glass, swim, boats, wet, waves

Second-Graders

Airplane: wings, fly, small airplane, air, helicopter, big, bird, land (verb), sky, pilot

Bed: sheets, pillow, mattress, sleep, sofa, blanket, quilt, soft, wood, bunk bed

Book: letters, sheets, read, write, note-book, school, stories, pencil, bookshop, pages

Car: wheels, steering-wheel, van, doors, truck, drive, window, seats, engine, motorbike

Clothing: sweater, pants, beautiful, dressing, dress, wash, wardrobe, coat, warm, shirt

Dog: puppy, cat, cute, bark, animal, doghouse, fleas, feet, ears, hair

Door: wood, gate, lock, hard, entrance, big, open, knob, window, keys

Face: eyes, mouth, nose, skin, body, ears, hair, happy, (a synonym of face), cheeks

Mother: father, son, loving, beautiful, tenderness, love, darling, hair, aunt, grandmother

Rain: water, umbrella, drops, clouds, cold, (a synonym of drops), wind, wet, raincoat, sky

Song: sing, microphone, music, voice, singer, loudspeakers, letter, listen, radio, poetry

Stone: hard, floor, rock, round, heavy, grey, wall, big, land, sculptor

Street: road, houses, cars, sidewalk, people, garden, floor, narrow, city, walk

Tooth: mouth, gum, denture, milk, white, rotten, dentist, tongue, cavities, bite

Tree: log, leaves, fruits, flowers, root, apples, sticks, branches, grass, garden

Water: fountain, river, rain, sea, drink, juice, tap, polluted, glass, bath

Preadolescents

Airplane: fly, air, sky, wings, trip, height, small airplane, jet, airport, clouds, travel, transport

Bed: sleep, sheets, pillow, mattress, blanket, comfortable, room, sleepy, dreams, lay, chair, rest

Book: read, pages, letters, school, study, reading, stories, sheets, front-page, pen, pencil, magazine

Car: wheels, engine, steering-wheel, van, door, ride, accident, drive, street, Ferrari, tires, transport

Clothing: pants, shirt, dress (verb), sweater, skirt, coat, dress, shorts, fashion, t-shirt, fabric, (a synonym of clothing)

Dog: she-dog, cat, animal, bark, leash, collar, friend, bite, kennel, hair, doghouse, tail

Door: knob, lock, house, wood, key, gate, entrance, open, bell, window, close, latch

Face: eyes, nose, mouth, (a synonym of face), (another synonym of face), ear, wash, beautiful, person, head, skin, smile

Mother: father, love, loving, family, darling, son, friend, good, beautiful, person, grandmother, tenderness

Rain: water, umbrella, drops, cloud, wet, weather, cold, thunderstorm, winter, sun, liquid, wind

Song: music, sing, microphone, voice, letter, singer, melody, flute, guitar, to rock, listen, radio

Stone: hard, rock, floor, land, nature, throw, big stone, big, granite, sidewalk, mason, heavy

Street: sidewalk, road, cars, houses, walk (verb), pavement, avenue, lamps, people, tar, quarter, floor

Tooth: cavity, dentist, mouth, white, canine, gum, chew, rotten, toothpaste, toothbrush, braces, wash

Tree: leaves, fruits, nature, log, root, flowers, forest, wood, plant, shade, green, branch

Water: drink, rain, river, wet, liquid, sea, fishes, thirst, tap, glass, fresh, pure

Adults

Airplane: trip, fly, air, sky, clouds, height, fear, transport, wings, freedom, quick, holidays, big, towers, terrorism

Bed: sleep, rest, sleepy, sheets, sex, warm, comfort, room, blankets, soft, love, (a synonym of rest), lay, death-bed, pillow

Book: read, culture, study, sheets, wisdom, history, knowledge, letters, pages, note-book, school, pleasure, romance, paper, learn

Car: transport, speed, wheels, motorbike, automobile, drive, road, ride, trip, vehicle, (a synonym of drive), Mercedes, red, Peugeot, gas

Clothing: dressing, pants, warm, fashion, coat, comfort, shopping, shops, cold, protection, sweater, underwear, vanity, beautiful, (warm clothes)

Dog: friend, cat, animal, companion, bark, cute, hair, puppy, loyal, fear, bone, collar, tender, beautiful, teeth

Door: entrance, house, window, opened, exit, open, wood, key, closed, lock, pathway, obstacle, knob, safety, bell

Face: (a synonym of face), eyes, beautiful, (another synonym of face), tails, person, beauty, expression, half, smile, round, ugly, mirror, happy, mouth

Mother: love, father, friend, loving, everything, beautiful, protection, darling, family, comfort, tenderness, happiness, life, missing, unique

Rain: water, cold, winter, wet, sun, sadness, umbrella, weather, storm, clouds, drops, discomfort, acid, wind, humidity

Song: music, joy, melody, cradle, sing, letter, sound, dance, Christmas, relaxation, (a synonym of song), voice, amusement, festival, love

Stone: hard, rock, mountain, cold, floor, (big stone), grey, philosopher, death, street, heavy, land, ridge of mountains, river, house

Street: road, house, sidewalk, path, cars, movement, go out, city, address, narrow, my, persons, pavement, exit, people

Tooth: mouth, white, dentist, pain, braces, eat, crunch, chew, smile, gum, cavities, toothbrush, bite, tongue, toothpaste

Tree: fruits, leaves, green, life, shade, nature, garden, oxygen, air, big, forest, country-side, branches, log, bird

Water: thirst, drink, life, sea, purity, fresh, limpid, transparent, river, wine, glass, blue, crystalline, swim, cold

Lists of Experiment 2

Lists Specific to Preschoolers

Book: read, see, stories, paint, drawings, to count, note-book, pages, [play, sheets, write, front-page]

Dog: ruff ruff, bark, bite, feet, lick, puppy, bones, ears, [eat, cat, hair, tale]

Door: open, close, key, enter, house, lock, white, wood, [leave, knob, go away, carpet]

Rain: water, wet, fall, drops, umbrella, sky, clouds, cold, [floor, winter, ill, coat]

Lists Specific to Preadolescents

Book: read, pages, letters, school, study, reading, stories, sheets, [front-page, pen, pencil, magazine]

Dog: puppy, cat, animal, bark, leash, collar, friend, bite, [kennel, hair, doghouse, tail]

Door: knob, lock, house, wood, key, gate, entrance, open, [bell, window, close, latch]

Rain: water, umbrella, drops, cloud, wet, weather, cold, thunderstorm, [winter, sun, liquid, wind]

Appendix B

Percentages of correct and false recall for each list and age

	Preschoolers		Second-graders		Preadolescents		Adults	
	Correct	False	Correct	False	Correct	False	Correct	False
Water	29.3	12.5	45.3	37.5	55.5	50.0	60.4	28.1
Tree	39.1	6.3	48.1	21.9	54.7	31.3	59.0	40.6
Airplane	30.1	3.1	50.6	3.1	58.1	18.8	64.0	15.6
Bed	32.4	3.1	50.9	34.4	63.0	43.8	65.0	21.9
Song	21.9	0.0	47.8	0.0	58.3	6.3	60.8	6.2
Dog	19.5	21.9	51.9	34.4	59.6	18.8	63.1	15.6
Face	37.9	3.1	63.4	3.1	67.5	6.3	56.9	21.9
Car	41.4	12.5	47.2	18.8	61.2	21.9	68.8	15.6
Rain	34.4	31.3	44.1	31.3	63.8	37.5	54.8	50.0
Tooth	23.4	3.1	52.8	46.9	68.0	37.5	63.5	12.5
Book	28.9	18.8	43.8	43.8	61.7	37.5	55.8	25.0
Mother	21.9	0.0	56.6	12.5	65.6	9.4	60.2	0.0
Stone	21.1	0.0	39.7	25.0	58.1	34.4	59.0	46.9
Door	29.7	15.6	42.8	25.0	59.6	37.5	60.6	40.6
Clothes	44.1	3.1	51.3	3.1	70.3	0.0	58.1	34.4
Street	42.2	3.1	46.3	3.1	67.7	15.6	56.0	15.6