The influence of theme identifiability on false memories: Evidence for age-dependent opposite effects

PAULA CARNEIRO

Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal

ANGEL FERNANDEZ

Universidad de Salamanca, Salamanca, Spain

AND

ANA RITA DIAS

Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal

In the present study, we used the Deese/Roediger–McDermott paradigm to analyze the relationship between theme identifiability of word lists and false memories in adults and children. We conducted two normative studies to determine the identifiability levels for critical unrepresented words in 40 associative lists in adults and in 16 associative lists in children. Then, in three experiments, false memories for critical words that were either easy or hard to identify were analyzed in adults and in children 4–5 years old and 11–12 years old. Opposite results were found for adults and children. Lists with highly identifiable critical words produced fewer false memories for adults but more false memories for children. These results suggest that, if they can identify critical words, adults use an identify-to-reject strategy to edit out false memories, whereas, in children, theme identification does not lead to the use of such a monitoring strategy.

The understanding of false memory phenomena frequently requires consideration of two opposite processes: an *error-inflating process* that results from the activation of distorted or nonexperienced information and an *error-editing process* that is responsible for the eventual rejection of a false memory (Arndt & Gould, 2006). It is the balance between those processes that determines whether a false memory is produced.

This dual process has been exemplified in one of the most powerful laboratory-based procedures for studying false memories: the DRM (after Deese, 1959; Roediger & McDermott, 1995) paradigm (see also Read, 1996). In this paradigm, lists composed of associates, but omitting the converging word (*critical word*), are presented to participants who are subsequently asked to perform a recall or recognition task, or both. Numerous experiments using this paradigm have demonstrated that rates of recall or recognition for the unrepresented critical words occur at high rates that often are similar to the rates of recall or recognition for the presented words. In most current views, two opposite processes combine to produce a false memory under these circumstances. In a first stage, the presentation of the list of associates could activate their unrepresented critical word, producing an error-inflating process. If this stage is not followed by an error-editing

process that allows for the rejection of the product of this error-inflating process, a false memory occurs (see Gallo, 2006, for an extensive review of major findings, consistent replications of the effects, and prevailing lines of explanation).

The error-inflating process has been denominated differently by the two main theories concerned with the explanation of the false memory illusion. In the activation-monitoring framework (Roediger, Balota, & Watson, 2001), it is called *activation*, whereas as part of fuzzy-trace theory (Brainerd & Reyna, 1998; Reyna & Brainerd, 1995), it is called *gist encoding*. Regardless of the differences in how these concepts are defined, both theories agree that this is usually the first step toward producing a false memory. The creation of a false memory occurs only if the unrepresented information has previously been activated or if the gist or the general meaning of the given information has previously been generated.

With respect to the error-editing processes, these two dual-process theories concur that not every activation or gist representation is necessarily manifest as a false memory, but they also present some differences in the way that this editing process is conceived. According to fuzzy-trace theory, the retrieval of true verbatim traces, or memory for specific information, can neutralize the familiarity of

false but gist-consistent events, leading to the suppression of false memories, an error-editing process termed *recollection rejection* (Brainerd & Reyna, 2002; Brainerd, Reyna, Wright, & Mojardin, 2003). However, within the activation-monitoring framework, this kind of process is referred to as *monitoring* (Roediger, Balota, & Watson, 2001), in accordance with Johnson's earlier concept of source monitoring (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981). According to the source-monitoring framework, the source is interconnected with all the specific conditions under which a memory has been encoded, which include both the perceptual attributes (i.e., sensory, temporal, and spatial information) and the cognitive operations involved in encoding. Within this framework, a false memory occurs if the source of an error-inflated memory is confused with the sources of true memories. Thus, a false memory could be rejected if one could distinguish the sources of the true and the error-inflated memories and could thus attribute the origin of a false memory to an internal generation process instead of to an actual external world event.

More recently, in order to emphasize two decision processes underlying false memory editing, Gallo (2006) distinguished two types of monitoring processes: *diagnostic* and *disqualifying*. Diagnostic monitoring refers to the rejection of an unrepresented item on the basis of the absence of an expected recollection, which generates the assumption that "if that item had occurred, I would have remembered it." This type of monitoring can occur as an automatic process or it can be more strategic, resembling the use of the distinctiveness heuristic (see Schacter, Israel, & Racine, 1999). Within the DRM paradigm, diagnostic monitoring would tend to occur if the critical words were highly memorable because of such characteristics as their emotionality (e.g., Pesta, Murphy, & Sanders, 2001) and word length (e.g., Madigan & Neuse, 2004) or because of their presentation format at the time of their study (e.g., pictures in Israel & Schacter, 1997).

Disqualifying monitoring refers to the rejection of an unrepresented event or item on the basis of a true recollection of another event or item, with the two being mutually exclusive. This type of monitoring, termed a *recall-to-reject process*, implies a strategic or conscious decision process. Recently, recall-to-reject processes have been studied extensively, mostly by proponents of dual-process models of recognition (see Yonelinas, 2002, for a review). From this perspective, such processes facilitate rejection of unstudied information via detection of features present in foil items that clearly indicate that the items are distractors. For example, a foil item could be rejected because it is remembered that it was presented in a different list (Rotello & Heit, 2000) or in a different modality (McElree, Dolan, & Jacoby, 1999).

In the DRM paradigm, one way to disqualify the critical lures is to know how many items were presented in each list and exhaustively recall all the items studied. Although effective in disqualifying the critical word, this strategy requires the correct recall of all the presented items, which is not usually what happens. Another way to use disqualifying monitoring to edit out false memories is to apply

the identify-to-reject strategy. Because of the converging nature of the lists, participants were able to notice the absence of the theme word from the list at the time of study and could use this information to reject a false memory at the time of the test.

Several types of evidence show that adults sometimes use the identify-to-reject strategy to edit out false memories in the DRM paradigm. Direct evidence comes from the responses to open-ended questionnaires filled in at the end of the DRM experiment, in which some participants freely report using this strategy to suppress false memories. In a study by Gallo, Roberts, and Seamon (1997), this strategy was the one most commonly reported (by 43.75% of participants), in cases in which participants had been informed about the false memory effect. Using a think-out-loud procedure, Lampinen, Meier, Arnal, and Leding (2005) also showed that participants sometimes self-report using similar strategies in order to reject critical items (e.g., by specifically remembering one word that was in fact presented: "Trash was not presented, because I remember *garbage*").

Stronger evidence that the application of this strategy decreases the level of false memories is provided by studies that encourage participants to use it in order to edit out their false memories. These studies give instructions that warn participants about the false memory effect and advise them to identify the critical words. For example, Neuschatz, Benoit, and Payne (2003) advised half of the participants not to make the typical DRM errors by showing them a list example. They were further instructed to figure out what the critical word of each list was and then note whether that word was in the presented list. In this experiment, the critical word was presented in some lists, and participants were informed that it would reappear from time to time throughout the experiment. Also, after each list presentation, participants were informed of the critical word. The results of this experiment showed that this type of instruction decreased false recognition for lists in which their critical words were more easily identified. A similar effect was also shown for a recall task in a study by Libby and Neisser (2001). In the second experiment of this study, not only were participants advised to identify the critical words, they were also instructed how to do this in order to suppress false recall. Participants were warned to identify an unrepresented "secret word" that was related to all the list words (by using the first list words to guess for the secret word) and then to confirm that this word was not in the list by paying attention to the remaining list words. The use of this strategy proved to be very efficient in eliminating false recall for trials in which participants guessed the secret word.

However, it is still uncertain whether participants spontaneously use this strategy if they are not advised about its effect or if they are not explicitly encouraged to search for the critical word. As argued above, the absence of a false memory could mean either that the critical word was not activated or that participants successfully suppressed the recall or recognition of the critical word. A study by Brédart (2000) suggests that the latter explanation is more plausible. In this study, each recall task was followed by

a second phase in which participants were asked to say all the words they had thought about and not recalled because they believed they were not presented in the list. If participants did not recall the critical words in the first phase, the probability of mentioning them in the second phase was significantly higher than the probability of not recalling them. This result means that the nonoccurrence of false memories is more likely to be caused by a strategy of the identify-to-reject type than by the nonactivation of the critical lure or the impossibility of getting the right gist of the list. Furthermore, a study by Mukai (2005) has shown that awareness of the nature of the list organization is related to a decrease in the level of false memories in an immediate recall test, and to an increase in the level of critical words produced in a postrecall test similar to the task used in the Brédart study.

The results of Brédart (2000) and Mukai (2005) seem to reinforce the idea that, at least sometimes, without being forewarned of the false memory effect, participants realize the nature of the list's organization and can spontaneously reject false memories through the use of the identify-to-reject strategy. However, we predict that identification of the theme of a list is one of the essential factors in originating this kind of spontaneous awareness and, consequently, in eliciting this editing process. More specifically, we predict that, even if participants are not forewarned, the correct identification of the critical word, and, as a result, the awareness that the critical word has not been presented in the list, will be greater for lists that have more identifiable themes than for lists that do not.

Following up on this idea, we tested false memories for lists that had both more and less identifiable critical words in a standard DRM paradigm. If theme identification is a crucial variable for the editing process, it is expected that lists that have more identifiable critical words will facilitate the editing process, generating fewer instances of false recall or false recognition. In this way, it would be possible to analyze whether participants spontaneously apply the identify-to-reject strategy in a standard DRM paradigm.

Such a prediction is well founded for young adults; however, in children, the interplay between error-inflating and -editing processes is a more controversial issue. It has been demonstrated that children can also show the false memory effect with the DRM paradigm, although to a lesser extent than can adults (Brainerd, Reyna, & Forrest, 2002; Carneiro, Albuquerque, Fernandez, & Esteves, 2007; Holliday, Reyna, & Brainerd, 2008; Holliday & Weekes, 2006; Howe, 2005, 2006; Howe, Cicchetti, Toth, & Cerrito, 2004; Howe, Gagnon, & Thouas, 2008; Lampinen, Leding, Reed, & Odegard, 2006; Odegard, Holliday, Brainerd, & Reyna, 2008). However, there has been mixed evidence on the issue of whether young children can efficiently use error-editing processes to reject false memories, with only a few studies providing some clues about this question. Brainerd, Holliday, and Reyna (2004) and Brainerd and Reyna (2002) provided a mathematical model of recognition to estimate the exact amounts of false memory editing that occur in children of different age levels. These authors found that, although false

memory editing increases with age, significant levels can be detected in young children. Furthermore, a study by Ghetti, Qin, and Goodman (2002) showed that children, like adults, reduce false memories through the presentation of distinctive information. This study showed that children 5 and 7 years of age made significantly fewer commission errors in a word-plus-picture encoding condition than in a word-encoding condition. This result suggests that distinctive information can increase the monitoring processes in all ages, which could mean that even young children are able to reject false memories in the DRM paradigm if the presented words are supported by distinctive information such as pictures.

On the other hand, a study by Rybash and Hrubí-Bopp (2000) showed that if the recall task was preceded by a generation task, children experienced source memory difficulties in a DRM paradigm. Through the inclusion of a phase in which participants were asked to think of words related to the study list immediately after the list presentation, children were more likely to falsely recall the critical words, potentially because of source-monitoring difficulties. But striking evidence suggests that, under some circumstances, children are even more capable than adults of suppressing false memories. In a study by Howe (2005), in which DRM lists were presented with directed-forgetting instructions, children, unlike adults, succeeded in decreasing their instances of false recall. According to Howe (2005), children's false memories are less automatic and are produced with more effort than are the false memories of adults. Thus, whereas for adults the production of false memories is automatic and likely to escape conscious control, for children, false memories imply a conscious processing that results in a facilitation of rejection.

In summary, the evidence stemming from available DRM studies cannot yet provide a clear answer with regard to whether or not young children can use error editing to reject false memories. As far as we know, disqualifying monitoring, or, more precisely, the identify-to-reject strategy, has not yet been studied in children, although there is some evidence regarding the use of diagnostic monitoring in children. Studies using a list-learning paradigm (Ghetti, 2003) or autobiographical memory (Ghetti & Alexander, 2004) indicate that young children do not make use of a memorability-based strategy, even if they are instructed to do so (Ghetti & Castelli, 2006). These studies show that, until the age of 9, children are not consistent in rejecting high- as opposed to low-memorability false events, which suggests that they do not apply the rule "If that had happened, I would have remembered it." Because this rule, based on the idea of expected memorability to infer event nonoccurrence, probably requires some cognitive control, it is particularly difficult for children to apply it. It is possible that other types of mechanisms, such as recollection rejection (Brainerd & Reyna, 2002), may be observed in young children. Actually, Brainerd and Reyna (2002) argued that, although there is an improvement in the verbatim memory throughout development, young children are sometimes able to spontaneously use verbatim traces of true events to suppress memories of false but gist-consistent events.

However, as has been established by Gallo (2006), identify-to-reject monitoring requires rather complex strategic processing, which is possibly not yet developed in young children. It is well known in developmental psychology that, if tasks require reliance on memory strategies, children's performance is impaired (Schneider & Bjorklund, 1998). The ability to spontaneously apply the kinds of retrieval strategies that are required in order to conduct appropriate memory searches emerges only in early adolescence. The identify-to-reject strategy, within the DRM paradigm, may be such a case, in that it requires the following steps: noticing that all the words in a list are related to a common theme, identifying the theme word that connects all the presented words, noticing that this theme word was not presented in the list, and holding it in mind in order not to recall it in the future. Although this kind of strategy may be rather easy for adults, it can be very difficult for children.

In order to clarify this question, in this study we also tested for children's false memories of lists that had both more and less identifiable critical words. If children are efficient in the spontaneous use of the identify-to-reject strategy, then high-identifiable lists should lead to lower levels of false memory. Otherwise, if children have difficulties in applying such a strategy, high-identifiable lists should promote gist extraction, and, in the absence of an efficient editing process, false memories should increase.

The goal of the present set of studies was to examine whether the identify-to-reject strategy is spontaneously used and whether the relationship between identifiability and false memories is the same for adults and children. Five studies were conducted: two normative studies aimed at empirically determining the degree of theme identifiability of a set of word lists, both in adults (Normative Study 1) and in children (Normative Study 2); and three experimental studies to examine the production of false memories with high- and low-identifiable lists in adults (Experiments 1 and 3) and in children (Experiment 2).

NORMATIVE STUDY 1

Because testing the aforementioned predictions requires manipulation of associative lists that are more or less likely to have an easy-to-identify critical word, the aim of this particular study was to obtain the relevant normative data for a sufficiently large set of lists with a sample of young adults. In a simple task, participants were presented with lists of words and were asked to generate, for each list, a single word that best defined its overall theme.

Method

Participants. One hundred ninety-four native speakers of Portuguese (117 women, 77 men; mean age, 27.3 years, $SD = 7.3$) were recruited from several undergraduate and postgraduate courses at Lisbon universities: social education ($n = 47$), tourism ($n = 33$), law ($n = 20$), psychology ($n = 25$), economics ($n = 44$), and accounting ($n = 25$). Because of the large number of lists (40), participants were divided into two similar groups to be tested with only half of the lists.

Materials. Forty Portuguese associative lists were selected from a pool of 100 lists already tested for false memories in a Portuguese-

speaking sample (Albuquerque, 2005). As in the original normative study, each list consisted of 15 associates, arranged in decreasing order of associative relationship to a critical word that was omitted at the time of presentation. The 40 lists embraced the full range (0%–78%) of false recall obtained in the previous study. The lists were divided into two sets that were similar in false recall probabilities. For the 20 lists in Set 1, false recall ranged from 4% to 78% ($M = 33.35$); for the 20 lists in Set 2, false recall ranged from 0% to 68% ($M = 30.55$). The words in each list were recorded on audio tape at a rate of 1.5 sec/word.

Procedure. Participants were tested in groups of up to 25 people. Approximately half of the participants were tested with 20 lists and the other half with the other 20 lists. In all groups, the instructions were identical:

You are going to listen to lists of related words, and for each list you must find the theme of the list. Please write the word that you think best describes the theme of each list. In case you are not sure, you have two more spaces to write three possible responses in total. For each response, please rate your level of confidence that your responses would be the theme of the list (on a 5-point scale from "not confident at all" to "extremely confident").

With the aim of not influencing participants' responses, these instructions did not include any example or make any reference to the inclusion or not of the theme word within the list.

After the oral presentation of each list, participants had approximately 30 sec to respond and to rate their confidence in relation to the correctness of each response. The session lasted about 20 min for each group of participants.

Results and Discussion

All the responses were recorded in a database and were kept in their original form, with the exception that singular/plural and masculine/feminine forms of the same word were combined into the form given most frequently by the participants as a whole. The results were analyzed, first by taking into account only the first response generated for each list, and then by considering the three possible responses.

Table 1 shows the number of participants who responded to each list, the number of different themes per list, the percentage of participants who indicated the critical word as the theme of the list, the English translation of the second theme generated, its production percentage, whether this second theme was included in the list, the English translation of the third theme generated (if equal to or above 4%, corresponding to 3 or more participants), its production percentage, and whether this third theme was included in the list.

The degree of identification of the critical word as the theme of the list differed considerably among different lists (ranging from 1% to 77%). The theme most frequently generated almost always coincided with the critical word. However, in 25% of cases (10 lists out of 40), the most frequently produced theme was not the critical word, but a word included in the presented list (this happened in 6 lists out of 10) or another word not presented (this happened in 4 lists out of 10). Taking into consideration only the themes generated for each list that were different from the critical words, it can be seen that 22 out of 40 of the second themes, and about 20 out of 33 of the third themes, were words from the presented lists, which means that about half of the second themes and

Table 1
Identifiability Indicators for the 40 Adult Lists of Normative Study 1

English Translation	Portuguese Critical Words	N	No. of Themes	% Critical Theme	Theme 2	% Theme 2	In List?	Theme 3	% Theme 3	In List?
Airplane	Avião	99	30	59.6	September 11th	4.0	n			
Bird	Pássaro	98	21	51.6	Freedom	10.1	y	Fowl	7.1	y
Butterfly	Borboleta	95	22	44.2	Spring	13.7	y	Bee	10.5	n
Chair	Cadeira	99	41	20.2	Furniture	7.0	n	Rest	7.0	y
Cinema	Cinema	95	20	66.3	Leisure	6.3	y	Recreation	4.1	y
Clown	Palhaço	99	20	64.7	Circus	6.1	y	Animation	5.1	n
Cold	Frio	99	14	10.1	Season	30.3	n	Winter	17.2	y
Cow	Vaca	95	17	76.8	Cattle	3.2	n			
Doctor	Doutor	93	29	10.5	Physician	30.5	y	Health	9.5	y
Dog	Cão	94	14	34.6	Animal	12.6	y	Friendship	5.3	n
Door	Porta	99	25	33.3	Habitation	26.3	n	House	12.1	y
Drug	Droga	99	21	44.4	Addiction	15.2	n	Vicious	9.0	y
Earth	Terra	94	21	28.4	Universe	17.9	y	Planet	13.7	y
Eyes	Olhos	95	20	61.7	Face	7.0	n	Vision	4.0	y
Floor	Chão	90	44	38.9	Earth	5.3	y			
Foot	Pé	94	32	31.5	Sport	9.5	n	Body	6.4	y
Garden	Jardim	98	29	51.5	Country	5.1	n	Paradise	5.1	n
Heart	Coração	98	14	70.7	Love	8.1	n	Passion	5.1	y
High	Alto	96	31	12.1	Height	32.3	n			
Hour	Hora	95	31	15.8	Weather	13.7	y	Punctuality/Routine	8.4	y/n
Illness	Doença	95	26	55.8	Health	5.3	y	Suffering	4.3	y
King	Rei	93	34	23.2	Power	13.7	y	Monarchy	11.6	y
Kiss	Beijo	95	20	4.2	Feelings	24.2	n	Sex/Love	13.7	n/y
Lion	Leão	97	16	73.7	Animal	8.1	y			
Meat	Carne	94	30	20.0	Feeding	13.7	n	Food	12.6	y
Music	Música	99	26	64.6	Amusement	4.0	n			
Painting	Quadro	99	23	19.2	Art	34.4	y	Painting	12.1	y
Pen	Caneta	95	35	1.1	Study	16.9	y	School	10.5	y
River	Rio	98	30	38.4	Water	13.1	y	Nature	7.1	n
Rubber	Borracha	94	42	21.1	School	6.3	y	Material	4.3	n
Rubbish	Lixo	99	20	51.5	Pollution	9.1	y	Rubbish bin	7.1	n
Sleep	Sono	95	25	17.9	Rest	24.2	y	Sleep	16.8	y
Spider	Aranha	98	19	72.7	Fear	5.0	y			
Stone	Pedra	95	40	40.4	Hardness	5.1	n	Mountain	4.0	y
Sun	Sol	99	27	17.2	Summer	20.2	y	Holidays	9.1	y
Sweet	Doce	95	25	25.3	Cooking	13.7	n	Food	12.6	n
Tie	Gravata	95	41	36.8	Suit	5.3	y	Formality	4.3	y
Wheel	Roda	94	43	12.6	Transport	15.8	n	Movement	8.4	y
White	Branco	93	40	14.1	Color	22.3	y	Sky	9.1	y
Window	Janela	90	43	20.0	Freedom	8.4	n	Nature	6.3	n

about a third of the third themes were words that were neither presented nor critical.

Table 2 displays the number of participants that gave a second and a third response to each list and the total percentage of identifiability of the critical word, taking into account the sum of all the three responses given for each list. As can be seen, the number of second responses was low (between 8 and 39, depending on the list, with an average of 19.28), and third responses were almost nonexistent (between 0 and 12, with an average of 3.53). Probably for this reason, the level of identifiability of the critical word, when one considers the three responses, did not differ considerably from the level of identifiability obtained when one takes into account only the first response (Pearson correlation between the two measures, $r = .99$, $p < .001$). This result seems to indicate that the analysis of the first response is sufficient for the present purposes, rendering the second and third responses unnecessary for the theme-identification task.

In order to understand whether identifiability of the critical word is related to false memories and to other factors taken into consideration in the regression analysis conducted by Roediger, Watson, McDermott, and Gallo (2001), a correlation analysis between identifiability (based on the first response), correct and false recall, correct and false recognition (obtained for the same lists used in the study of Albuquerque, 2005), forward association,¹ frequency, and length of critical items was conducted for the 40 lists (displayed in Table 3). A negative correlation was found between identifiability and false recall ($r = -.44$, $p < .01$), suggesting that high identifiability is reliably related to low false recall. No other variables correlated with identifiability.

To analyze the themes' confidence judgments, a 2 (type of theme: critical vs. other) $\times$ 2 (type of lists: more vs. less identifiable) ANOVA was performed, with repeated measures over the type of theme. We selected the 10 most and 10 least identifiable lists and then, on the basis of the

Table 2
Normative Study 1: Number of Participants That Generated the Critical Word in Their First Response (N 1st), in Their Second Response (N 2nd), and in Their Third Response (N 3rd), and Identifiability Percentages Taking Into Account All Three Responses (%I 1+2+3) or Only the First Response (%I 1)

Lists	N 1st	N 2nd	N 3rd	%I 1+2+3	%I 1
Airplane	99	10	3	59.6	59.6
Bird	98	11	0	55.6	51.6
Butterfly	95	21	3	48.5	44.3
Chair	99	20	3	21.2	20.2
Cinema	95	21	6	69.5	66.3
Clown	99	11	3	67.7	64.7
Cold	99	31	6	14.1	10.1
Cow	95	14	3	78.9	76.8
Doctor	93	22	3	11.6	10.5
Dog	94	22	5	39.9	34.6
Door	99	15	1	38.4	33.3
Drug	99	17	4	46.4	44.4
Earth	94	26	5	31.5	28.4
Eyes	95	11	0	64.7	61.7
Floor	90	14	5	42.1	38.9
Foot	94	18	3	33.6	31.5
Garden	98	14	2	53.5	51.5
Heart	98	13	3	71.7	70.7
High	96	16	2	13.1	12.1
Hour	95	28	5	16.9	15.8
Illness	95	25	5	68.5	55.8
King	93	25	8	28.5	23.2
Kiss	95	26	7	6.3	4.2
Lion	97	8	0	74.7	73.7
Meat	94	25	4	22.1	20.0
Music	99	17	3	67.6	64.6
Painting	99	18	2	19.2	19.2
Pen	95	32	8	1.1	1.1
River	98	21	1	41.4	38.4
Rubber	94	17	3	22.2	21.1
Rubbish	99	19	0	53.5	51.5
Sleep	95	22	5	20.0	17.9
Spider	98	13	1	75.7	72.7
Stone	95	11	1	41.4	40.4
Sun	99	16	1	18.2	17.2
Sweet	95	39	12	38.0	25.3
Tie	95	26	4	40.0	36.8
Wheel	94	26	7	12.6	12.6
White	93	11	0	14.1	14.1
Window	90	19	4	21.1	20.0

first response, recorded the average confidence ratings for the critical theme and for the other themes provided for those lists. The analysis revealed main effects of type of theme [$F(1,18) = 80.34$, $MS_e = .13$, $p < .001$] and type of list [$F(1,18) = 11.27$, $MS_e = .11$, $p < .01$], with an

interaction between the two [$F(1,18) = 9.17$, $MS_e = .13$, $p < .01$]. These results showed that, in general, the critical theme produced higher levels of confidence than did the other themes ($M = 4.03$ vs. $M = 3.00$) and that this difference was higher for more identifiable lists (an increment of 1.37) than for less identifiable lists (an increment of 0.68). For the other themes, the rates of confidence were very similar for more and less identifiable lists ($M = 2.997$ and $M = 3.000$, respectively), and the difference between them was not significant (post hoc Bonferroni analysis, $p > .05$). It is possible to conclude that participants were more confident of choosing the right theme if they generated the critical theme, especially with more identifiable lists.

EXPERIMENT 1

The previous normative study provided theme identifiability adult norms for 40 DRM lists, thus allowing for the manipulation of this variable in Experiment 1. To study the effect of theme identifiability on false memories, we selected the lists with the highest and the lowest levels of critical word identifiability.

As reported above, there was a significant correlation between the identifiability levels of the 40 lists and the false recall obtained in a previous study for the same lists (Albuquerque, 2005), in the sense that higher critical word identifiability is related to lower false recall. Because this analysis was performed with data obtained from different samples, at different times and places, and because it could suggest only that these variables were related, we designed the present study to investigate a causal relation between those variables within the same sample and under controlled experimental conditions. On the basis of the results of the correlational analysis, as well as the theoretical assumptions considered in the introduction, it was predicted that the lists with the highest levels of critical word identifiability would produce lower levels of false memories.

Method

Participants. The participants were 40 young adult volunteers, 26 females and 14 males, 20 to 40 years of age ($M = 26.08$, $SD = 4.83$). They were young professionals (57.5%) and undergraduate students (43.5%). All were native Portuguese speakers, and none had participated in the previous normative study.

Materials. Twelve Portuguese associative lists were selected, according to the identifiability scores obtained in Normative Study 1. Six lists had the highest scores for critical word identifiability: *cinema*, *clown*, *lion*, *cow*, *heart*, and *spider* (64.7%–76.8%). The other 6 had the lowest levels of identifiability: *cold*, *doctor*, *high*, *kiss*, *pen*, and *wheel* (1.1%–12.6%). High- and low-identifiable lists did not differ significantly in the number of letters of their critical words ($M = 5.67$ vs. $M = 4.83$, respectively), in their frequency ($M = 861.00$ vs. $M = 755.67$), or in their forward association ($M = 81.83$ vs. $M = 80.14$). Each list comprised the 15 strongest associates and was prerecorded at a rate of 1.5 sec/word. The lists were presented in a random order, so that they were different for each participant.

The recognition test comprised 60 words: 24 presented words (the 1st and the 8th word of each presented list) and 36 unpresented words (12 critical words of the presented lists, 12 unpresented related words [the 16th associate of the presented lists], and 12 unrelated words [drawn from other lists]). The words were disposed in five random orders, and each version was presented to 8 participants.

Table 3
Correlation Between Theme Identification and Correct and False Recall, Correct and False Recognition, Forward Association, Frequency, and Length

Factor	Theme Identification
Correct recall ^a	.07
False recall ^a	-.44**
Correct recognition ^a	.21
False recognition ^a	-.16
Forward association ^b	-.13
Frequency ^c	-.07
Length	.18

^aData from Albuquerque (2005). ^bData from Albuquerque (2001). ^cData from Linguistic Center of Lisbon University. ** $p < .01$.

Procedure. The experiment was conducted in small groups of up to 5 participants. The instructions for the recall and recognition tasks followed standard DRM instructions (see, e.g., Roediger & McDermott, 1995). The participants were initially informed that they would be performing a memory task. Before presentation of the lists, the participants were informed that they would hear lists of words played on a tape recorder for a later recall task. After the presentation of each list, participants were given 2 min to write down as many words as they could remember. Immediately after the recall of the last list, the participants were given a self-paced final recognition task, in which they were instructed to read the words written on a sheet of paper and to decide whether these words had been presented earlier. For each word, they had to circle the “yes” answer if they thought that the word had been presented and the “no” answer if they thought it had not. The participants were also instructed to rate their confidence in their responses (from *not confident at all* to *extremely confident*, on a 5-point scale). An analysis of recollective experience can provide a way of checking whether participants are engaging in an editing process in order to reject false memories (Lampinen, Watkins, & Odegard, 2006). Participants applying the identify-to-reject strategy in order to reject false memories of highly identified critical items should show a tendency to express high confidence in their rejection of those words.

On both tests, the participants were instructed not to guess. Each experimental session lasted approximately 35 min.

Results and Discussion

Table 4 displays the recall and recognition data: the proportions of studied and critical words recalled and the proportions of studied words, critical words, and related distractors recognized for both types of lists.

Recall. True and false recall were analyzed by a 2 (item type: studied vs. critical) $\times$ 2 (identifiability: high vs. low) ANOVA, in which both factors were within subjects. The results showed a main effect of type of item [$F(1,39) = 71.88$, $MS_e = .02$, $p < .001$], revealing that true recall ($M = .58$) was higher than false recall ($M = .33$). There was also a main effect of identifiability [$F(1,39) = 13.30$, $MS_e = .02$, $p < .01$], reflecting that low-identifiable lists ($M = .49$) produced higher levels of recall than did high-identifiable lists ($M = .42$). Finally, there was a significant item $\times$ identifiability interaction [$F(1,39) = 13.43$, $MS_e = .02$, $p < .01$], clarifying the overall pattern of the recall data. Post hoc Bonferroni analyses confirmed that the difference between low- and high-identifiable lists was nonsignificant for correct recall ($p > .05$), but statistically significant for false recall ($p < .01$). Although identifiability did not have an effect on correct recall, it significantly influenced false recall, with low-identifiable lists producing higher levels

of false recall than did high-identifiable lists. The more plausible explanation for this result is that the difficulty in extracting the theme from the low-identifiable lists did not permit the use of the identify-to-reject strategy. And, because of this, the critical word that was activated but not identified could not be so effectively rejected as being an unrepresented word.

Recognition. The recognition data included only 39 participants, because 1 participant missed several responses. Overall, participants showed good performances, with a high proportion of correct recognition ($M = .79$) and a very low proportion of false alarms to unrelated distractors ($M = .04$). True and false recognition were analyzed by a 3 (item type: studied vs. critical vs. related distractor) $\times$ 2 (identifiability: high vs. low) ANOVA, with both factors being within subjects. There was a main effect of type of item [$F(2,76) = 141.18$, $MS_e = .05$, $p < .001$], in which studied items ($M = .79$) produced higher levels of recognition than did critical items ($M = .50$) and unrelated distractors ($M = .21$). Pairwise comparisons (Bonferroni analysis) showed that the differences among the three types of items were significant (all $ps < .001$). In addition, a main effect of identifiability [$F(1,38) = 14.38$, $MS_e = .02$, $p < .01$] showed that low-identifiable lists ($M = .54$) produced higher recognition than did high-identifiable lists ($M = .47$). There was also a significant item type $\times$ identifiability interaction [$F(2,78) = 9.29$, $MS_e = .03$, $p < .001$]. Post hoc analyses showed that the difference in recognition between high- and low-identifiable lists was significant only for the critical items ($p < .001$). The difference in identifiability was not significant for either studied items or related distractors ($p > .05$). This general pattern of recognition corroborates the results of the recall task. Both tasks provided evidence leading to the conclusion that list identifiability determines only false memories and not true memories. However, it is known that if the recognition test follows a recall task, a contamination of the recognition results can occur from the recall task (Roediger & McDermott, 1995). Thus, in accordance with the recall data, recognition results should be interpreted with caution.

In order to find additional evidence for the claim that an editing process was responsible for the lower level of false memories of the high-identifiable lists, we separated the confidence ratings for critical words that were rejected (“no” responses) and for those that were falsely recognized (“yes” responses) for each list. Then, a 2 (response type: rejection vs. false recognition) $\times$ 2 (identifiability: high vs. low) ANOVA was performed, with the latter factor within subjects. The results showed a main effect of type of response and a significant response type $\times$ identifiability interaction. The main effect of response type [$F(1,10) = 13.27$, $MS_e = .08$, $p < .01$] revealed that critical words that were falsely recognized ($M = 4.31$) produced higher rates of confidence than did critical items that were rejected ($M = 3.90$). The significant item type $\times$ identifiability interaction [$F(1,10) = 16.56$, $MS_e = .08$, $p < .01$], together with planned comparisons, confirmed that the rejection confidence was significantly higher for high-identifiable critical items ($M = 4.29$) than for low-identifiable critical items ($M = 3.51$) ($p < .01$). However,

Table 4
Proportions of Different Types of Responses for Low- and High-Identifiable Lists (LI, HI) on the Recall and Recognition Tasks of Experiment 1

	LI Lists		HI Lists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recall				
Studied	.57	.07	.58	.05
Critical	.41	.21	.25	.24
Recognition				
Studied	.78	.12	.81	.12
Critical	.60	.23	.41	.31
Related distractors	.24	.21	.18	.19

no significant difference was found between high ($M = 4.25$) and low ($M = 4.38$) identifiability for critical items that were falsely recognized. These results clearly indicated that participants were more confident that the critical words were not presented if those words were easily identified, which suggests that an editing process was in operation for rejecting high-identifiable critical items.

NORMATIVE STUDY 2

Like Normative Study 1, this study was designed to generate identifiability norms for DRM lists, but this time for a child population. Because lists specially tailored to the respective age of children have been shown to be more fruitful for generating false memories (Carneiro et al., 2007), in order to analyze their identifiability levels, we used 16 associative lists derived from specific associative norms for different ages. In this way, we could be certain that all the words were from children's vocabularies and that they contained age-specific associates. We could further establish relations with the same lists already tested for false memories. In this study, we selected two age groups, preschoolers and preadolescents, because these are sufficiently distant from one another in their course of development to show any eventual differences in the relationship between identifiability and false memories.

Method

Participants. Seventy-five children participated in this study: 50 preschoolers (27 females, 23 males) 4 and 5 years of age ($M = 4$ years, 11 months) and 25 preadolescents (14 females and 11 males) 11 and 12 years of age ($M = 11$ years, 10 months). All were native Portuguese speakers recruited from the kindergartens and middle schools of a small city on the west coast of Portugal.

Materials. Sixteen Portuguese associative lists specially assembled for children were constructed on the basis of free association norms for 16 Portuguese words given to children of different ages (3–4, 7–8, and 11–12) in a previous study (Carneiro, Albuquerque, Fernandez, & Esteves, 2004). The critical words were common nouns within the vocabulary of young children, each with only one

meaning. The specific lists for preschoolers and preadolescents each contained the 10 strongest associates, arranged in a decreasing order and omitting the critical word. The lists were previously recorded on an audio tape at a rate of 1 word every 1.5 sec. For preschoolers, the lists were divided into two blocks of eight lists each.

Procedure. Preadolescents were tested in small groups (up to 8 participants in each) with the 16 lists, and written responses were required. Because younger children are, in general, more affected by the length of the test, preschoolers were tested individually with 8 lists and were asked to provide oral responses. In order to test all the 16 lists, two groups of preschoolers were formed, with 25 participants each.

The participants were instructed to listen to the lists of related words and to find the theme for each list. After listening to each list, the preschoolers verbalized the theme of the list and the preadolescents wrote it down on an appropriate sheet of paper. The instructions were adapted to the children's age. For the preschoolers, it took the form of a game and required additional explanations.

Overall, the procedure of this experiment was the same as that used for adults in Normative Study 1. However, two main changes were introduced in order to adapt it to a more simplified version for young children. First, just one response was requested for each list (the results of the normative study for adults showed that the percentage of additional responses was very low, and it did not add any relevant information to the analyses). Furthermore, confidence ratings for the theme word were not included, because it was thought that this would be a difficult task for young children.

The lists were presented in a random order, different for each participant. The session lasted approximately 15 min.

Results and Discussion

As in the adult normative study, words that differed only in number or grammatical gender were grouped together.

Table 5 shows the number of different themes per list, the percentage of participants who indicated the critical word as the theme of the list, and the second theme generated and its percentage (if it is at least 12%, corresponding to 3 or more participants). It also shows whether this second theme had been in the list, for both age groups. As in the case of the adults tested with a larger number of lists, the identifiability level of the critical word in both age groups differed among the 16 lists (4%–92% for preschoolers,

Table 5
Identifiability Indicators for the 16 Lists With Preschoolers and Preadolescents in Normative Study 2

Lists	Preschoolers					Preadolescents				
	No. of Themes	% Critical Theme	Theme 2	% Theme 2	In List?	No. of Themes	% Critical Theme	Theme 2	% Theme 2	In List?
Airplane	7	52	Bird	24	n	9	64			
Bed	5	68	Sleep	20	y	7	60	Rest	12	n
Book	10	44	Draw	12	n	8	72			
Car	5	84				5	80			
Clothing	12	36	Dressing	16	y	5	48	Garments	40	n
Dog	3	92				5	60	Animals	28	y
Door	5	64	House	24	y	6	72	House	12	y
Face	10	32	Head	28	y	8	56	Person	12	y
Mother	9	60				9	28	Family	32	y
Rain	5	76				8	48	Weather	16	y
Song	16	4	Sing	20	y	14	20	Band	20	n
Stone	9	32	Ball	36	n	12	52			
Street	14	4	Car	20	y	14	36	City	12	n
Tooth	5	80				6	76			
Tree	10	52	Flowers	12	y	10	56			
Water	9	40	Beach	16	y	9	68			

20%–80% for preadolescents). In the majority of lists, the most frequently generated theme corresponded to the critical word. In one list for preadolescents and three for preschoolers, the themes most frequently produced were words that were different from the critical word.

To analyze the ability to extract the theme from the lists, a paired *t* test was performed, comparing the identifiability levels of the 16 lists for preschoolers and preadolescents. This comparison showed no significant differences between the two ages (for preschoolers, $M = 51.25$, $SD = 26.22$; for preadolescents, $M = 56.00$, $SD = 17.03$). This result seems to contradict the assumption that young children have special difficulties in extracting the gist of the lists (Brainerd et al., 2002). On average, there were no significant differences, but preschoolers made a greater distinction between high- and low-identifiable lists than preadolescents. It seems that the critical words (e.g., *dog*, *car*) of some lists were very well identified by preschoolers, whereas others (e.g., *song*, *street*) were identified by just 1 participant of this age group.

Following the same line of analysis conducted for adults, correlations between identifiability levels and other potentially relevant factors were computed for the 16 lists (see Table 6). For each age group, the proportions of correct recall, false recall, correct recognition, and false recognition (reported in a previous study by Carneiro et al. [2007] for the same ages and lists) were selected in order to analyze their relationships with identifiability. Other variables, such as forward association² and the frequency and length of critical items, were also included in the analysis. The results revealed significant positive correlations between the identifiability level of each age group and false recall ($r = .52$, $p < .05$, for preschoolers; and $r = .58$, $p < .05$, for preadolescents). Furthermore, a positive correlation was found between identifiability and false recognition in preadolescents ($r = .51$, $p < .05$). And, surprisingly, we also observed positive correlations between identifiability and the length of critical items ($r = .60$, $p < .05$, for

preschoolers; and $r = .63$, $p < .01$, for preadolescents). Critical words composed of more letters were more easily identified as the themes of the lists.

In contrast with the results for adults (Normative Study 1), high-identifiable lists were those that had shown higher levels of false recall in a previous study (Carneiro et al., 2007). This result suggests that the relation between identifiability and false recall for adults is opposite to that for children.

EXPERIMENT 2

The results of Normative Study 2 were used to construct appropriate materials for the present experiment. In line with the methodology of Experiment 1, the lists with the highest and lowest levels of identifiability for each age group were selected to be directly tested on false memories.

As mentioned in the introduction, there is mixed evidence regarding the ability of children to suppress false memories in the DRM paradigm. Besides this, the identify-to-reject strategy has never been directly studied in children. Thus, taking into account the results of the aforementioned correlational analyses between identifiability levels and false recall, we predicted that lists with the highest levels of identifiability would produce higher levels of false memories for both age groups than would lists with low identifiability.

Method

Participants. Sixty children, 30 preschoolers (16 females, 14 males) 4 and 5 years of age ($M = 5$ years, 0 months) and 30 preadolescents (17 females, 13 males) 11 and 12 years of age ($M = 12$ years, 1 month), were recruited from the kindergartens and middle schools of a small city on the west coast of Portugal. All were native Portuguese speakers and none had participated in the normative study.

Materials. Eight associative Portuguese lists were selected on the basis of the identifiability scores obtained in Normative Study 2. The four most and the four least identifiable lists were selected for both age groups. In the preschool group, the most identifiable lists were those corresponding to the critical words *dog*, *car*, *tooth*, and *rain* (identifiability levels [IL], 76%–92%), and the least identifiable lists corresponded to the critical words *song*, *street*, *face*, and *stone* (IL, 4%–32%). For preadolescents, the most identifiable lists corresponded to the critical words *car*, *tooth*, *book*, and *door* (IL, 72%–80%) and the least identifiable ones corresponded to the critical words *song*, *street*, *mother*, and *clothing* (IL, 20%–48%). For both age groups, high- and low-identifiable lists did not differ significantly in the length of their critical words (preschoolers, $M = 4.5$ and $M = 4.5$; preadolescents, $M = 5.0$ and $M = 4.25$), in frequency (preschoolers, $M = 1,770.75$ and $M = 2,304.50$; preadolescents, $M = 3,288.25$ and $M = 2,543.25$), or in their forward association (preschoolers, $M = 71.67$ and $M = 70.23$; preadolescents, $M = 78.45$ and $M = 75.75$).

Eight lists were presented to each age group, each containing the 10 strongest associates. In some lists, 1 associate was omitted because it coincided with the critical word of another list. In these cases, the conflicting associate was replaced with the 11th associate, which was placed in the last position on the list. The lists were pre-recorded for presentation at a rate of 1 word every 1.5 sec, and they were presented randomly, in a different order for each participant.

The recognition test included 33 words: 16 presented words, corresponding to the 1st and 5th word of each list; 8 critical words; and

Table 6
Correlation Between Theme Identification of Preschoolers and Preadolescents and Correct and False Recall, Correct and False Recognition, Forward Association, Frequency, and Length

	Theme Identification	
	Preschoolers (PS)	Preadolescents (PA)
Theme identification PS	1.00	
Theme identification PA	.59*	1.00
Correct recall PS ^a	-.17	.12
Correct recall PA ^a	-.04	-.17
False recall PS ^a	.52*	.36
False recall PA ^a	.36	.58*
Correct recognition PS ^a	-.06	.17
Correct recognition PA ^a	.29	.31
False recognition PS ^a	-.10	-.02
False recognition PA ^a	.00	.51*
Forward association PS ^b	-.05	.29
Forward association PA ^b	.15	-.13
Frequency ^c	-.18	.15
Length	.60*	.63**

^aData from Carneiro et al. (2007). ^bData from Carneiro et al. (2004)

^cData from Linguistic Center of Lisbon University. * $p < .05$. ** $p < .01$.

9 unrelated words, corresponding to the critical, 1st, and 5th word of three unrepresented lists. The words in the recognition test were presented in a random order, different for each participant.

Procedure. This experiment followed the procedure used for adults in Experiment 1, with some adaptations for children, as was the case in the study of Carneiro et al. (2007): Participants were tested individually and their responses were given orally, the instructions were adapted to the children's respective ages, and the recognition task did not include confidence ratings.

Initially, the participants were informed that they would be performing a memory test (instruction for preadolescents) or a memory game (instruction for preschoolers), which consisted of listening to words played on a tape recorder and recalling them later. For each age group, the presentation order of the lists was counterbalanced across participants. They were allowed to spend 2 min recalling each list. After the recall of all lists, they performed a self-paced final recognition task, for which they were instructed to pay attention to the words, orally presented by the experimenter, and to decide whether each word had been presented on the tape recording. The words were presented in a different random order for each participant. For both recall and recognition tasks, the participants' responses were recorded by the experimenter on a score sheet.

Results and Discussion

Table 7 shows the recall and recognition data for both age groups.

Recall. True and false recall were analyzed by a 2 (item type: studied vs. critical) $\times$ 2 (age group: preschoolers vs. preadolescents) $\times$ 2 (identifiability: high vs. low) ANOVA, with repeated measures over item type and identifiability. The results showed three main effects: item type [$F(1,58) = 261.32$, $MS_e = .02$, $p < .001$], in which true recall ($M = .47$) was higher than false recall ($M = .15$); age group [$F(1,58) = 42.62$, $MS_e = .03$, $p < .001$], in which preadolescents ($M = .38$) showed higher levels of recall than did preschoolers ($M = .24$); and identifiability [$F(1,58) = 42.47$, $MS_e = .02$, $p < .001$], in which high-identifiable lists ($M = .37$) produced higher levels of recall than did low-identifiable lists ($M = .25$). There was also a significant age $\times$ identifiability interaction [$F(1,58) = 6.10$, $MS_e = .02$, $p < .05$], showing that the difference in recall between high- and low-identifiable lists was higher for preadolescents than for preschoolers. Post hoc Bonferroni analyses confirmed that the difference between these two types of lists was significant for both ages ($p < .01$ for preschoolers, $p < .001$ for preadolescents). Furthermore, a significant item type $\times$ identifiability interaction [$F(1,58) = 28.32$, $MS_e = .02$, $p < .001$] indicated that, in

contrast to the results for true recall, high-identifiable lists produced a considerably higher level of false recall than did low-identifiable lists. Post hoc analyses confirmed that the difference in false recall between high- and low-identifiable lists was significant ($p < .001$). However, although the difference in true recall reached acceptable significance ($p < .05$), it corresponded to a minimal difference between high- and low-identifiable lists. And finally, a significant item type $\times$ age $\times$ identifiability interaction [$F(1,58) = 6.79$, $MS_e = .02$, $p < .05$] emerged, confirming that for both ages, the difference in true recall between high- and low-identifiable lists was lower than the difference in false recall between the two types of lists (especially in the case of preadolescents).

Several conclusions can be drawn from this analysis. First, younger children produced considerably lower levels of true and false memories than did older children. The significant difference found between preschoolers and preadolescents regarding true and false recall is consistent with previous research that has analyzed the development of false memories with the DRM paradigm (for a review, see Brainerd, Reyna, & Ceci, 2008). Second, as far as the main prediction of this study is concerned, high-identifiable lists produced higher levels of false recall for both age groups. This means that, for children, theme identification results in more instances of false recall. Unlike adults, children cannot use gist extraction information to reject false memories, which indicates that they are not able to apply the identify-to-reject strategy to edit out false memories. And, third, although the difference in false recall between less and more identifiable lists was higher for preadolescents, it corresponds to a fourfold increase in both age groups. For this reason, we can conclude that, as in the case of preschoolers, preadolescents are not yet able to use gist identification to suppress false memories. Instead, for both preschoolers and preadolescents, gist identification seemed to increase false recall.

Recognition. An initial comparison of false alarms between preschoolers and preadolescents did not show a significant difference ($M = .03$ and $M = .04$, respectively).

As in free recall, recognition data were also analyzed by a 2 (item type: studied vs. critical) $\times$ 2 (age group: preschoolers vs. preadolescents) $\times$ 2 (identifiability: high vs. low) ANOVA, with repeated measures over type of item and identifiability. And, as in the case of the results of the recall task, three main effects emerged: item type [$F(1,58) = 95.92$, $MS_e = .04$, $p < .001$], in which true recognition ($M = .57$) was higher than false recognition ($M = .31$); age group [$F(1,58) = 43.77$, $MS_e = .06$, $p < .001$], in which preadolescents ($M = .55$) recognized more items than did preschoolers ($M = .34$); and identifiability [$F(1,58) = 47.05$, $MS_e = .04$, $p < .001$], in which high-identifiable lists ($M = .53$) produced higher levels of recognition than did low-identifiable lists ($M = .35$). In addition, there was also a significant item type $\times$ identifiability interaction [$F(1,58) = 11.84$, $MS_e = .03$, $p < .01$], indicating that, although the difference between high- and low-identifiable lists was significant for both true and false recognition (post hoc analyses, $p <$

Table 7
Proportions of Different Types of Preschoolers' and Preadolescents' Responses for Low- and High-Identifiable Lists (LI, HI) on the Recall and Recognition Tasks in Experiment 2

	Preschoolers				Preadolescents			
	LI Lists		HI Lists		LI Lists		HI Lists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recall								
Studied	.38	.09	.40	.10	.54	.08	.55	.07
Critical	.03	.10	.14	.17	.08	.13	.37	.30
Recognition								
Studied	.44	.17	.52	.20	.60	.18	.73	.18
Critical	.10	.18	.29	.20	.28	.26	.58	.26

.001), it was higher for false recognition than for true recognition.

Thus, in regard to age differences, preadolescents recognized more studied and critical items than did preschoolers, confirming the results of previous research (e.g., Brainerd et al., 2002; Carneiro et al., 2007). With respect to the identifiability effect, high-identifiable lists presented higher levels of correct and false recognition than did low-identifiable lists. These results are in line with the results of the recall task and corroborate the idea that gist extraction is fundamental for the formation of false memories in children.

EXPERIMENT 3

Since, in the previous experiments, children and adults were tested with associative lists that differed in their critical words, we decided to conduct one more experiment in which adults were tested with the same critical words as those used in Experiment 2 for children. This seemed to be especially relevant because, as was previously mentioned, the 40 lists for adults and the 16 lists for children could not be entirely controlled for all variables that are known to affect false memories (e.g., backward associative strength).

Method

Participants. The participants were 41 undergraduate students (29 females, 12 males), 20–45 years of age ($M = 30.92$, $SD = 8.63$). All were native Portuguese speakers and none had participated in the previous studies.

Materials and Procedure. Fifteen of the 16 associative critical words used in Experiment 2 had already been tested for adults by Albuquerque (2005) in regard to false recall and false recognition.³ The associative lists of 4 of those critical words were included in the identifiability norms for adults of the present study, and hence their levels of theme identifiability were already known. For the remaining 11 lists, a complementary theme-identifiability normative study was carried out prior to this experiment. Using the same instructions as in Normative Study 1, 92 participants were asked to generate a single word that best characterized the overall theme of those 11 lists.

The lists were selected from this pool of 15 lists that shared the critical words already tested for children in Experiment 2. The identifiability level of the critical words differed among the 15 lists, ranging from 33.3% to 91.86%, and it did not correlate with the levels of false recall and false recognition produced in the Albuquerque (2005) study. As in Experiment 2, 8 lists (the 4 most and the 4 least identifiable lists) were selected. The most identifiable lists (IL, 85.39%–91.86%) and the least identifiable lists (IL, 33.3%–54.95%) did not significantly differ in the length of their critical words, in frequency or in forward association.

The participants in this experiment were tested with these eight lists, each containing the 15 strongest associates (the same number of associates presented in the theme-identifiability normative study). This experiment followed the same general procedure as Experiment 2, except that the participants were tested in small groups (up to 14 persons) and they provided written responses.

The recognition task included 48 words: 16 presented words (the 1st and the 8th word of each presented list) and 32 unpresented words (the 8 critical words of the presented lists and 24 unrelated words from another eight unpresented lists [the critical words, and the 1st and the 8th associates]).

Results and Discussion

Table 8 displays the recall and recognition data for both types of lists.

Table 8
Proportions of Different Types of Adults' Responses for Low- and High-Identifiable Lists (LI, HI) on the Recall and Recognition Tasks of Experiment 3

	LI Lists		HI Lists	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recall				
Studied	.53	.08	.53	.08
Critical	.43	.29	.27	.26
Recognition				
Studied	.73	.16	.86	.13
Critical	.78	.23	.66	.30

Recall. True and false recall were analyzed by a 2 (item type: studied vs. critical) $\times$ 2 (identifiability: high vs. low) ANOVA, in which both factors were within subjects. Post hoc Bonferroni analyses ($p < .05$) were performed on all significant interactions between the independent variables. The results showed a main effect of item type [$F(1,40) = 24.62$, $MS_e = .03$, $p < .001$], revealing that true recall ($M = .53$) was higher than false recall ($M = .35$); a main effect of identifiability [$F(1,40) = 10.92$, $MS_e = .03$, $p < .01$], indicating that low-identifiable lists ($M = .48$) produced higher levels of recall than did high-identifiable lists ($M = .40$); and a significant item type $\times$ identifiability interaction [$F(1,40) = 9.53$, $MS_e = .02$, $p < .01$], clarifying that low-identifiable lists produced higher levels of false recall than did high-identifiable lists ($M = .43$ vs. $M = .27$) but similar levels of correct recall.

Recognition. One participant was excluded from the recognition analyses due to incomplete responses. In general, responses showed a high proportion of correct recognition ($M = .79$) and a low proportion of false alarms ($M = .04$).

True and false recognition were analyzed using the same factorial design as the one performed for the recall data. The results showed a significant item type $\times$ identifiability interaction [$F(1,39) = 36.71$, $MS_e = .02$, $p < .001$], revealing that high-identifiable lists produced higher levels of correct recognition ($p < .001$) but lower levels of false recognition ($p < .01$) than did low-identifiable lists. Furthermore, a marginal effect of item type [$F(1,39) = 4.07$, $MS_e = .02$, $p = .05$] showed that correct recognition ($M = .79$) tended to be higher than false recognition ($M = .72$).

Overall, these results confirmed that, for adults, false recall and false recognition decreased if critical items were easily identified. For lists that shared the same critical words, opposite results were found for adults and children in both recall and recognition tasks.

GENERAL DISCUSSION

The results of these experiments demonstrate two main findings. First, if the theme of the lists is easily identified, adults reduce their level of false memories. This study indicates that this happens even if participants are not warned about the false memory effect. For this reason, one can conclude that adults may spontaneously use the identify-to-reject strategy to edit out false memories in the DRM para-

digm. Because adult participants could easily identify the critical word, they could apply this strategy to reject false memories efficiently. It is, however, important to notice that these results did not match the null results obtained in the unwarned condition by Neuschatz et al. (2003) (a possible explanation for this discrepancy is considered below). Moreover, the result observed in Experiment 1 regarding a higher rejection confidence for high-identifiable lists seems to support the interpretation that, in fact, participants are using the identify-to-reject strategy to reject false memories of highly identified critical items.

Second, children until 12 years of age do not tend to spontaneously use the identify-to-reject strategy. The role of list identifiability differed according to the participants' age: Lists with more identifiable critical words caused lower levels of false memories for adults but higher levels of false memories for children. This suggests that ease in identifying the critical word can lead to a higher rejection of false memories in adults and to the production of more false memories in children.

A reasonable conclusion from this pattern of results is that, whereas adults can easily apply the identify-to-reject strategy, children have difficulties in using it. It is well known that, in general, young children demonstrate poor source-monitoring skills mainly if the memories are self-generated (Foley & Johnson, 1985; Foley, Johnson, & Raye, 1983) or if the sources are very similar (Lindsay, Johnson, & Kwon, 1991). This type of evidence revealed that, in some circumstances, young children are more prone to source memory confusions, which are thought to increase suggestibility (e.g., Ackil & Zaragoza, 1995; Lindsay, 2002; Perner, 2000). Because the identify-to-reject strategy is one of the monitoring processes that can be used to efficiently reject false memories, it would be predictable that an early age group, such as preschoolers, would show some difficulties in applying this strategy.

With respect to preadolescent children, recent research shows that, in general, they are already apt to use monitoring strategies such as the memorability-based strategy (Ghetti, 2003; Ghetti & Alexander, 2004; Ghetti & Castelli, 2006), or to use verbatim traces to reject the critical lures at retrieval. The finding that, in the present study, preadolescents failed to use theme identification to reject false memories could lead to the conclusion that the identify-to-reject strategy requires more cognitive effort than do memorability-based strategies. In fact, the identify-to-reject strategy implies several insights that could be complex even for preadolescents. It seems plausible that preadolescents are able to perform the first two phases of the identify-to-reject strategy—namely, noticing that all the list words are related to a common theme and, most of the time, identifying the related theme word. However, it is possible that preadolescents are not capable of engaging in a systematic search that would allow them to understand that the theme word was not presented in the list. The present study showed only that the identify-to-reject strategy is difficult for children, but it could not provide information about which of the phases in the application of this strategy children fail to achieve. Probably

there are differences in the development of these phases between preschoolers and preadolescents. This would be an interesting matter to explore in future research.

The finding that high-identifiable lists increase false memories in children reveals that, to a certain extent, even young children are sensitive to gist manipulation. In general, previous studies have found an inverted U-shaped relation between age and gist cuing (Bjorklund, Miller, Coyle, & Slawinski, 1997). Brainerd et al. (2008) argued that gist manipulations seemed to be less effective if participants are too young to process semantic relations or if they already have fully developed semantic abilities, as is the case in young adulthood. In the DRM paradigm, a similar pattern was observed. The instruction to form meaning-based connections before the presentation of DRM lists was maximally effective in later childhood or early adolescence, specifically from 8–11 years of age (Brainerd et al., 2004; Holliday et al., 2008; Lampinen, Leding, et al., 2006; Odegard et al., 2008). For example, Lampinen, Leding, et al. (2006) showed that 6-year-olds did not benefit from gist instructions, whereas 9-year-olds did. But more important for this debate would be to discover whether even young children's false memories can be affected by gist manipulations. Although not totally consistent, there is cumulative evidence to suggest that kindergarten and 1st-year primary school children could increase false memories if they were tested with material that emphasized the overall theme or if they were instructed to form meaning-based connections. Dewhurst, Purslove, and Lewis (2007) showed that, if the gist extraction is favored by presenting the stimuli in a story context, false recognition in 5-year-olds was increased. Similarly, studies by Brainerd, Forrest, Karibian, and Reyna (2006) and Holliday et al. found higher levels of false recall for children of 6 years of age or above who received meaning connection instructions. On the other hand, Howe (2006) found no effects of meaning connection instructions for categorical material in the false recall of 5-, 7-, and 11-year-olds. More in accordance with Dewhurst et al. (2007), Brainerd et al. (2006), and Holliday et al., the present study clearly indicates that young children produce more false memories if they can extract the gist of the lists and that, unlike adults, they cannot use gist extraction to reject false memories, probably because they lack the monitoring strategies or they are less able to use verbatim traces.

Surprisingly, the present study also reveals that the ability of preschoolers to extract the gist of the lists was not considerably inferior to that of preadolescents. This could indicate that young children are able to extract the gist of the lists in an identification task, but that in usual memory tasks they are not as efficient as older children in spontaneously extracting the gist, which, in general, results in lower false memories. Another possibility is that the enhanced ability of gist extraction by younger children observed in the present study could be the result of the nature of the lists. As was previously mentioned, the associative lists of the present study were generated from free-associative norms provided by children of the same

age. As was shown in the study by Carneiro et al. (2007), it is likely that gist extraction is facilitated if lists are appropriate to children's age.

The results of this study could also have implications for current theoretical explanations of the false memory phenomenon. The finding that more identifiable lists produce fewer false memories in adults cannot be easily accommodated by theories that attribute the responsibility of false memory only to the extraction of the thematic representation of the list. If this were the case, more identifiable lists should produce more false memories. But, rather than an increase in false recall or recognition, we found a facilitation of the rejection process if the theme of the lists was identified more easily. Although fuzzy-trace theory conceives of an editing process that is responsible for false memory rejection, it would also predict that higher gist identification would lead to higher levels of false memories in adults. Furthermore, young children's ability to extract the gist of child-normed lists can be considered another finding that could be difficult to handle by this theory, in view of its assumptions concerning the development of gist-processing abilities. Regarding the activation-monitoring theory, the results fit very well with a monitoring account, exemplifying the functioning of one of its strategies (although it can also be viewed as recollection rejection) and accounting for the different results between adults and children. Still, the results show that gist identification, a concept that has usually been dismissed by the supporters of the activation-monitoring account, is an important variable that has a role in explaining false memories in children and false memory rejection in adults.

The normative studies also provided interesting results. Standard DRM lists are usually derived from normative studies that provide the associates and their associative relationship to the critical word. These lists have been thoroughly studied from the point of view of their level of association, identifying indexes of forward and backward association, and relating them to their levels of false recall and false recognition (Brainerd & Wright, 2005; Deese, 1959; Gallo & Roediger, 2002; McEvoy, Nelson, & Komatsu, 1999; Roediger, Watson, et al., 2001). Roediger, Watson, et al. (2001) provided an extensive set of regression analyses, showing that backward association is a powerful factor in predicting false memories and, therefore, in determining the lists' variability in producing this phenomenon. Furthermore, variables such as connectivity (Deese, 1959; McEvoy et al., 1999), concreteness, and production frequency in the language (Roediger, Watson, et al., 2001) have been examined as possible factors that could be related to higher or lower levels of false memories. On the basis of the results of the present normative studies, identifiability may be one more aspect to consider in understanding the persistent finding that, although lists are built with the same criteria, they are nonetheless quite different in terms of false memory production.

In both normative studies, associative lists differed in terms of how they allowed for the intentional identification of their theme or gist. List differences in identifi-

ability were found in different age groups of preschoolers, preadolescents, and adults. Even for children, it is possible to distinguish lists in which the themes are easier or harder to identify. Although the ranges of critical word identification differ across age groups (adults, 1%–77% for the 40 lists included in Normative Study 1 and 33%–92% for the 15 lists in Experiment 3; preadolescents, 20%–80%; preschoolers, 4%–92%), in all ages, there is substantial variability between lists in regard to identifiability.

One possible criticism of this study might be that attention was not paid to a crucial factor, backward associative strength, because of material limitations. Since there are no known studies that could provide mean backward associative strength (MBAS) indicators for Portuguese lists, it was not possible to control this variable. This kind of normative data would have allowed us to examine identifiability while holding MBAS constant between high- and low-identifiable lists. However, we do not think that this aspect diminishes the strength of the inferences drawn for two main reasons. First, in the adult literature, false recall has been positively correlated with the lists' MBAS (Deese, 1959; Hicks & Hancock, 2002; Robinson & Roediger, 1997; Roediger, Watson, et al., 2001), which leads us to expect that the same finding would hold for Portuguese lists. But because in the present study false recall was negatively correlated with identifiability in adults, we do not think that MBAS could be considered a third variable that could account for the results. It might be difficult to explain why identifiability would be negatively correlated to MBAS. Second, even supposing that lists could differ in terms of MBAS, opposite results for adults and children for lists that share the same critical words would also be difficult to explain by differences in backward associative strength. One way of explaining the discrepancy between the results of the present study and those reported by Neuschatz et al. (2003) could be related to this issue. It is possible that the high-identifiable lists employed by those authors were related more strongly to the critical lures than were the high-identifiable lists of the present study. Thus, if association strength and identifiability work in opposite ways to produce false memories, one could argue that, in the unwarned condition of Neuschatz et al., association strength and identifiability canceled each other and produced a null effect. However, the validity of this explanation, opposing associative strength and identifiability in false memory explanation, cannot be established without additional research.

In summary, the adult-related results of the present study highlight the role of one type of monitoring, the identify-to-reject strategy, to edit out false memories in the DRM paradigm, and they show that adults are likely to use this strategy spontaneously if the critical word is easily identifiable. In contrast, children are less able to use this strategy, and for this population, high identifiability tends to lead to more false memories. In short, it is reasonable to conclude that identifiability has different effects according to age: It can promote an error-editing process in adults and an error-inflating process in children.

AUTHOR NOTE

The first author was supported by Fundação para a Ciência e a Tecnologia (SFRH/BPD/26729/2006). A.F. was supported by the Spanish Ministry of Education and Science and the Ministry of Science and Innovation (Projects SEJ2005-03210 and PSI2008-05607, respectively) and by Junta de Castilla y León (Project SA031A-06). We acknowledge the comments by M. Cortese and by anonymous reviewers of earlier drafts, and we thank Raquel Mendes for her help in collecting children's data. Correspondence should be addressed to P. Carneiro, Universidade Lusófona de Humanidades e Tecnologias, Departamento de Psicologia, Av. do Campo Grande, 376, 1749-024, Lisboa, Portugal (e-mail: mpcarneiro@usal.es).

REFERENCES

- ACKIL, J. K., & ZARAGOZA, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology*, **60**, 57-83. doi:10.1006/jecp.1995.1031
- ALBUQUERQUE, P. B. (2001). *Normas de associação semântica de palavras portuguesas para aplicar ao paradigma DRM*. Relato Técnico: Universidade do Minho.
- ALBUQUERQUE, P. B. (2005). Produção de evocações e reconhecimentos falsos em 100 listas de palavras associadas portuguesas. *Laboratório de Psicologia*, **3**, 3-12.
- ARNDT, J., & GOULD, C. (2006). An examination of two-process theories of false recognition. *Memory*, **14**, 814-833. doi:10.1080/09658210600680749
- BJORKLUND, D. F., MILLER, P. H., COYLE, T. R., & SLAWINSKI, J. L. (1997). Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training studies. *Developmental Review*, **17**, 411-441. doi:10.1006/drev.1997.0440
- BRAINERD, C. J., FORREST, T. J., KARIBIAN, D., & REYNA, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology*, **42**, 962-979. doi:10.1037/0012-1649.42.5.962
- BRAINERD, C. J., HOLLIDAY, R. E., & REYNA, V. F. (2004). Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child Development*, **75**, 505-522. doi:10.1111/j.1467-8624.2004.00689.x
- BRAINERD, C. J., & REYNA, V. F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*, **71**, 81-129. doi:10.1006/jecp.1998.2464
- BRAINERD, C. J., & REYNA, V. F. (2002). Recollection rejection: How children edit their false memories. *Developmental Psychology*, **38**, 156-172. doi:10.1037/0012-1649.38.1.156
- BRAINERD, C. J., REYNA, V. F., & CECI, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*, **134**, 343-382. doi:10.1037/0033-2909.134.3.343
- BRAINERD, C. J., REYNA, V. F., & FORREST, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development*, **73**, 1363-1377. doi:10.1111/1467-8624.00477
- BRAINERD, C. J., REYNA, V. F., WRIGHT, R., & MOJARDIN, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, **110**, 762-784. doi:10.1037/0033-295X.110.4.762
- BRAINERD, C. J., & WRIGHT, R. (2005). Forward association, backward association, and the false-memory illusion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 554-567. doi:10.1037/0278-7393.31.3.554
- BREDART, S. (2000). When false memories do not occur: Not thinking of the lure or remembering that it was not heard? *Memory*, **8**, 123-128. doi:10.1080/096582100387669
- 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.
- CARNEIRO, P., ALBUQUERQUE, P., FERNANDEZ, A., & ESTEVES, F. (2007). Analyzing false memories in children with associative lists specific for their age. *Child Development*, **78**, 1171-1185. doi:10.1111/j.1467-8624.2007.01059.x
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22. doi:10.1037/h0046671
- DEWHURST, S. A., PURSGLOVE, R. C., & LEWIS, C. (2007). Story contexts increase susceptibility to the DRM illusion in 5-year-olds. *Developmental Science*, **10**, 374-378. doi:10.1111/j.1467-7687.2007.00592.x
- FOLEY, M. A., & JOHNSON, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development*, **56**, 1145-1155. doi:10.2307/1130229
- 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. doi:10.2307/1129860
- GALLO, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. New York: Psychology Press.
- GALLO, D. A., ROBERTS, M. J., & SEAMON, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, **4**, 271-276.
- GALLO, D. A., & ROEDIGER, H. L., III (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory & Language*, **47**, 469-497. doi:10.1016/S0749-596X(02)00013-X
- GHETTI, S. (2003). Memory for nonoccurrences: The role of metacognition. *Journal of Memory & Language*, **48**, 722-739. doi:10.1016/S0749-596X(03)00005-6
- GHETTI, S., & ALEXANDER, K. W. (2004). "If it happened, I would remember it": Strategic use of event memorability in the rejection of false autobiographical events. *Child Development*, **75**, 542-561. doi:10.1111/j.1467-8624.2004.00692.x
- GHETTI, S., & CASTELLI, P. (2006). Developmental differences in false-event rejection: Effects of memorability-based warnings. *Memory*, **14**, 762-776. doi:10.1080/09658210600648548
- GHETTI, S., QIN, J., & GOODMAN, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology*, **38**, 705-718. doi:10.1037/0012-1649.38.5.705
- HICKS, J. L., & HANCOCK, T. W. (2002). Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin & Review*, **9**, 807-815.
- HOLLIDAY, R. E., REYNA, V. F., & BRAINERD, C. J. (2008). Recall of details never experienced: Effects of age, repetition, and semantic cues. *Cognitive Development*, **23**, 67-78. doi:10.1016/j.cogdev.2007.05.002
- HOLLIDAY, R. E., & WEEKES, B. S. (2006). Dissociated developmental trajectories for semantic and phonological false memories. *Memory*, **14**, 624-636. doi:10.1080/09658210600736525
- HOWE, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science*, **16**, 927-931. doi:10.1111/j.1467-9280.2005.01638.x
- HOWE, M. L. (2006). Developmentally invariant dissociations in children's true and false memories: Not all relatedness is created equal. *Child Development*, **77**, 1112-1123. doi:10.1111/j.1467-8624.2006.00922.x
- HOWE, M. L., CICCETTI, D., TOT, S. L., & CERRITO, B. M. (2004). True and false memories in maltreated children. *Child Development*, **75**, 1402-1417. doi:10.1111/j.1467-8624.2004.00748.x
- HOWE, M. L., GAGNON, N., & THOUAS, L. (2008). Development of false memories in bilingual children and adults. *Journal of Memory & Language*, **58**, 669-681. doi:10.1016/j.jml.2007.09.001
- ISRAEL, L., & SCHACTER, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review*, **4**, 577-581.
- JOHNSON, M. K., HASHTRUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28. doi:10.1037/0033-2909.114.1.3
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, **88**, 67-85. doi:10.1037/0033-295X.88.1.67
- LAMPINEN, J. M., LEDING, J. K., REED, K. B., & ODEGARD, T. N. (2006). Global gist extraction in children and adults. *Memory*, **14**, 952-964. doi:10.1080/09658210601008957
- LAMPINEN, J. M., MEIER, C. R., ARNAL, J. D., & LEDING, J. K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 954-963. doi:10.1037/0278-7393.31.5.954
- LAMPINEN, J. M., WATKINS, K. N., & ODEGARD, T. N. (2006). Phantom ROC: Recollection rejection in a hybrid conjoint recognition

- tion signal detection model. *Memory*, **14**, 655-671. doi:10.1080/09658210600648431
- LIBBY, L. K., & NEISSER, U. (2001). Structure and strategy in the associative false memory paradigm. *Memory*, **9**, 145-163. doi:10.1080/741948851
- LINDSAY, D. S. (2002). Children's source monitoring. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony* (pp. 83-98). Chichester, U.K.: Wiley.
- LINDSAY, D. S., JOHNSON, M. K., & KWON, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, **52**, 297-318. doi:10.1016/0022-0965(91)90065-Z
- MADIGAN, S., & NEUSE, J. (2004). False recognition and word length: A reanalysis of Roediger, Watson, McDermott, and Gallo (2001) and some new data. *Psychonomic Bulletin & Review*, **11**, 567-573.
- MC ELREE, B., DOLAN, P. O., & JACOBY, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 563-582. doi:10.1037/0278-7393.25.3.563
- MC EVOY, C. L., NELSON, D. L., & KOMATSU, T. (1999). What is the connection between true and false memories? The differential roles of interitem associations in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1177-1194. doi:10.1037/0278-7393.25.5.1177
- MUKAI, A. (2005). Awareness of the false memory manipulation and false recall for people's names as critical lures in the Deese-Roediger-McDermott paradigm. *Perceptual & Motor Skills*, **101**, 546-560. doi:10.2466/PMS.101.6.546-560
- NEUSCHATZ, J. S., BENOIT, G. E., & PAYNE, D. G. (2003). Effective warnings in the Deese-Roediger-McDermott false-memory paradigm: The role of identifiability. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 35-41. doi:10.1037/0278-7393.29.1.35
- ODEGARD, T. N., HOLLIDAY, R. E., BRAINERD, C. J., & REYNA, V. F. (2008). Attention to global gist processing eliminates age effects in false memories. *Journal of Experimental Child Psychology*, **99**, 96-113. doi:10.1016/j.jecp.2007.08.007
- PERNER, J. (2000). Memory and theory of mind. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 297-312). Oxford: Oxford University Press.
- PESTA, B. J., MURPHY, M. D., & SANDERS, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 328-338. doi:10.1037/0278-7393.27.2.328
- READ, J. D. (1996). From a passing thought to a false memory in 2 minutes: Confusing real and illusory events. *Psychonomic Bulletin & Review*, **3**, 105-111.
- REYNA, V. F., & BRAINERD, C. J. (1995). Fuzzy-trace theory: An interitem synthesis. *Learning & Individual Differences*, **7**, 1-75. doi:10.1016/1041-6080(95)90031-4
- ROBINSON, K. J., & ROEDIGER, H. L., III (1997). Associative processes in false recall and false recognition. *Psychological Science*, **8**, 231-237. doi:10.1111/j.1467-9280.1997.tb00417.x
- ROEDIGER, H. L., III, BALOTA, D. H., & WATSON, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger III, 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. doi:10.1037/10394-006
- ROEDIGER, H. L., III, & McDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 803-814. doi:10.1037/0278-7393.21.4.803
- ROEDIGER, H. L., III, WATSON, J. M., McDERMOTT, K. B., & GALLO, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, **8**, 385-407.
- ROTELLO, C. M., & HEIT, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*, **28**, 907-922.
- RYBASH, J. M., & HRUBI-BOPP, K. L. (2000). Source monitoring and false recollection: A life span developmental perspective. *Experimental Aging Research*, **26**, 75-87. doi:10.1080/036107300243696
- SCHACTER, D. L., ISRAEL, L., & RACINE, C. (1999). Suppressing false recognition: The distinctiveness heuristic. *Journal of Memory & Language*, **40**, 1-24. doi:10.1006/jmla.1998.2611
- 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, & language* (5th ed., pp. 467-521). New York: Wiley.
- YONELINAS, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517. doi:10.1006/jmla.2002.2864

NOTES

1. *Forward association* refers to the strength of association between the critical lure and the list items when the critical lure serves as the cue. A related measure is *backward association*, which refers to the strength of such association when the list items serve as a cue. Backward association was not included in the analysis, because, unfortunately, there are no available data regarding this variable for Portuguese associative lists.
2. Backward association was not included, because, as for adults, there are no available data for children's Portuguese lists.
3. One of the critical words studied for children had not been included in the Albuquerque (2005) study.

(Manuscript received July 30, 2008;
revision accepted for publication September 28, 2008.)