False Memories in Children and Adults: Age, Distinctiveness, and Subjective Experience

Simona Ghetti
University of California, Davis

Jianjian Qin
Yale University

Gail S. Goodman
University of California, Davis

This study investigated developmental trends associated with the Deese/Roediger–McDermott false-memory effect, the role of distinctive information in false-memory formation, and participants’ subjective experience of true and false memories. Children (5- and 7-year-olds) and adults studied lists of semantically associated words. Half of the participants studied words alone, and half studied words accompanied by pictures. There were significant age differences in recall (5-year-olds evinced more false memories than did adults) but not in recognition of critical lures. Distinctive information reduced false memory for all age groups. Younger children provided with distinctive information, and older children and adults regardless of whether they viewed distinctive information, expressed higher levels of confidence in true than in false memories. Source attributions did not significantly differ between true and false memories. Implications for theories of false memory and memory development are discussed.

In the early 1990s, Henry Roediger (Roediger & McDermott, 1995) and Don Read (1996) independently revived an experimental paradigm originally designed by Deese (1959). They demonstrated that robust false-memory effects can be induced in adults by simply requiring them to remember lists of words. Specifically, in this paradigm, individuals study one list at a time of semantically related words. At the end of each list, individuals are then asked to recall the words they just studied. For example, participants are asked to remember a word list such as thread, pin, eye, sewing, sharp, point, pricked, thimble, haystack, pain, hurt, injections. When asked to recall these words, many participants report not only the studied words but also the word needle (referred to later in this article as the “critical lure”), which was not included in the list, thus demonstrating that a sizable percentage of adults may falsely recall nonpresented words if they are strong semantic associates of the presented words. These findings have been reliably observed with lists of semantically related words converging on a variety of themes. In addition, this false-memory effect is observed in recognition memory tests. When after studying several word lists, participants are presented with a recognition test that includes studied words, critical lures, and nonsemantically associated lures, participants are as likely to falsely recognize critical lures as they are to correctly recognize studied words. The probability of obtaining this false-memory effect in both recall and recognition is a function of the total associative strength of the list words, that is, of the sum of the probabilities that each word will elicit the critical lure for individuals involved in a free-association task.

Although intrusions of critical lures elicited in Deese’s (1959) paradigm had been known as “implicit associative responses” (Underwood, 1965) well before Roediger and McDermott (1995) and Read (1996) published their results, these latter authors recognized that when participants claimed to have heard stimuli that were never presented in the lists, they were experiencing false memories. This insight made Roediger and McDermott’s article a classic in the literature on false memory and spurred a deluge of research (e.g., Dodson & Schacter, 2001; Mather, Henkel, & Johnson, 1997; McDermott, 1996; Norman & Schacter, 1997; Payne, Elie, Blackwell, & Neuschatz, 1996; Robinson & Roediger, 1997).

In the last decade, there has been extraordinary interest in studies of false memory, catalyzed partly by legal cases on recovered memories of child sexual abuse and partly by a shift from a quantity-oriented to an accuracy-oriented conceptual framework in the study of memory (Koriat, Goldsmith, & Pansky, 2000). Because robust false-memory effects can be easily and reliably obtained using the Deese/Roediger–McDermott (DRM) paradigm and because this paradigm guarantees a level of experimental control that is often difficult to match when studying false-memory formation for real-world events, researchers have relied on the task to address several important questions concerning the nature of false memories in adults.
The DRM paradigm may also be important for the study of false-memory formation from a developmental perspective. Using other paradigms (e.g., misinformation), researchers have uncovered reliable age trends in false memory. However, it is unclear whether reliable age trends would emerge in the DRM task, which relies on semantically associated word lists. Although it has been demonstrated that both children and adults are more likely to falsely recognize items that are semantically associated with studied materials than items that are not, results regarding age are not always consistent (for a review, see Brainerd & Reyna, 1998). As discussed shortly, from a theoretical perspective, different predictions can be made regarding the type of age trends expected with the DRM paradigm. Studying developmental trends with the DRM paradigm was an essential goal of the present study.

In addition to questions about age trends, other important questions arise from former research that used the DRM paradigm. For example, do the same factors that enable adults to suppress the creation of false memory reduce false-memory formation in children? In this regard, one possible factor is the encoding of distinctive information. Presentation of distinctive information reduces false-memory formation in adults (e.g., Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999; Smith & Hunt, 1998). Whether it also reduces false-memory formation in children was examined in the present study. Furthermore, researchers have often questioned, for adults, whether false memories are different phenomenologically from true memories (e.g., Lampanen, Neutschat, & Payne, 1998; Qin, 1999). In the present study, we investigated age differences in the phenomenological experience of true and false memory.

In conducting developmental research with the DRM task, it should be noted that, because the DRM paradigm involves memorizing many words, it could be considered a demanding task for young children. Thus, researchers should ensure that child participants are given sufficient time to encode and retrieve the material. Further, the number of items to be included in each word list should be carefully considered so that it is adequate to elicit the false-memory effect but does not overwhelmingly exceed children’s memory abilities, such as their immediate memory span (Dempster, 1981). These factors were considered when the present study was designed.

Before explaining in detail the rationale of the present study, we should mention one special advantage of using the DRM paradigm in a developmental study of false memory. The social demands placed on participants by the paradigm are not as high as those found in several other false-memory paradigms (e.g., the misinformation paradigm or the “lost in the mall” paradigm; see McCloskey & Zaragoza, 1985). Thus, if participants report recalling or recognizing a word that was not in the studied list, researchers can be relatively confident that participants are actually experiencing a false memory rather than merely succumbing to social compliance.

The Developmental Trajectory of Spontaneous False Memories

The large body of research stemming from the DRM paradigm with adults suggests important lines of inquiry for developmental psychologists. One such line concerns the investigation of developmental changes in spontaneous false memories. On the one hand, it may be hypothesized that false-memory formation decreases with age. According to fuzzy-trace theory (FTT), individuals store in memory two separate representations of events, namely, gist traces and verbatim traces (for reviews, see Brainerd & Reyna, 1998; Brainerd, Reyna, & Poole, 2000; Reyna & Brainerd, 1995). Both gist traces, which preserve the meaning of the item, and verbatim traces, which hold contextual item-specific information, develop with age. In the context of the DRM paradigm, verbatim traces may prevent individuals from forming false memories by sustaining nonidentity judgments. These judgments are due to the feeling of contrast resulting from the inner comparison between verbatim traces of studied items and gist traces of nonstudied items. Older children have more enduring verbatim memory for studied material than do younger children, and they can better use the retrieved correct information to make nonidentity judgments, that is, to edit out similar but nonstudied distractors (e.g., Reyna & Kiernan, 1994). Thus, when such comparison is performed, older children may be less likely than younger children to endorse related distractors, which could potentially result in lower rates of false-memory formation for older children on the DRM task.

On the other hand, the prediction that false-memory effects may increase with age is also consistent with FTT as a consequence of developmental differences in the formation of gist traces. Beyond simple automatic activation, gist traces maintain the semantic meaning of the studied material and could elicit false memories by supporting judgments of similarity as well as phantom recollection, a vividly recollective state (Brainerd, Wright, Reyna, & Mojardin, 2001). Because the ability to extract gist from context is believed to increase with age, so should the likelihood of creating false memories as a result of studying word lists converging on a theme (i.e., gist). Thus, to the extent that false-memory effects with DRM lists are driven by spontaneous elaboration on the meaning of the list members, younger children may show a reduced false-memory effect for semantically related distractors. Some research supports this prediction, which contrasts with the typical finding that younger ages are associated with heightened memory distortion. For example, Bach and Underwood (1970) involved second and sixth graders in a word-recognition test in which acoustic and associative relations were varied. In their study, sixth graders were more likely than second graders to produce false alarms in response to associatively related distractors. Felzen and Anisfeld (1970) found similar results when comparing false recognition for semantically related items in third graders and sixth graders.

Opposite predictions regarding developmental trends may also be derived from Johnson’s source monitoring framework (SMF; Johnson, Hastroud, & Lindsay, 1993). Within the SMF, false memories are conceptualized as a failure to correctly attribute the source of activated information. According to the SMF, event features (e.g., semantic content of the material, visual and auditory characteristics of the stimulus, the experimental context, traces of cognitive processes) are bound together as a result of encoding processes. Subsequently, on any particular occasion, a cue is likely to activate some subset of those encoded features, which may vary in clarity and veridicality. Recognition and recall both involve identifying the source of the activated information, and false memory arises when internally derived information (i.e., informa-
tion that is generated through mechanisms such as associative activation of the semantic network) is erroneously attributed to an external source (e.g., having studied the item in a word list). Because young children are generally more susceptible to source-monitoring errors than are older children and adults (e.g., Foley & Johnson, 1985; Foley, Johnson, & Raye, 1983; Foley & Ratner, 1998; Lindsay, Johnson, & Kwon, 1991; see Quas, Schaaf, Alexander, & Goodman, 2000, for a review), young children should be more likely than adults to falsely attribute the automatic activation of the critical lures in the DRM paradigm to an external source and should come to believe that the critical lures were studied words. It should be noted, however, that this prediction will hold only if the effects of other processes (e.g., the strength of automatic associative activation during encoding and retrieval) are comparable across age groups. Some of these processes are likely to be age dependent and may work in the opposite direction of developing source-monitoring capability in the context of the DRM paradigm (e.g., the strength of the semantic association between *needle* and *injection* may be weaker in younger children than in older children and adults). It is therefore possible that younger children will evince less false memory despite underdeveloped source-monitoring capabilities because of limitations in the semantic association of critical lures with other members of the lists.

Thus, it is apparent that contrasting predictions for developmental changes in false memory with DRM lists may be advanced. Nevertheless, it should be noted that despite possible developmental differences in verbatim memory, decision processes, and the like (which could result in decreases in false-memory formation with age), because the DRM paradigm presents participants with word lists converging on a theme, the DRM task seems particularly tailored to emphasize gist extraction and relational processing among items. Thus, the DRM paradigm should offer, more than most other experimental procedures that include semantically associated material, appropriate conditions to observe age-related increases in false-memory formation. However, given the numerous mental processes that may contribute to performance on the DRM task, developmental data are needed to constrain theory and guide precise predictions.

The Role of Distinctiveness in the Reduction of False-Memory Formation

Psychologically, distinctiveness derives from the processing of differences relative to some context or background (Howe, Courage, Vernescu, & Hunt, 2000). Such processing may result in enhanced memory performance. With the DRM lists, the encoding of distinctive information along with words reduces false-memory formation in adults (Israel & Schacter, 1997; Read, 1996; Schacter et al., 1999; Smith & Hunt, 1998). Investigating whether this is true for children as well was one of the goals of the present study. To accomplish this goal, we operationalized distinctiveness as in Schacter et al.’s studies (Israel & Schacter, 1997; Schacter et al., 1999); that is, words were either studied alone or along with pictures representing them.

For adults, two mechanisms concerning distinctiveness have been proposed to account for the false-memory reduction phenomenon. Predictions about whether young children’s false memory will be reduced after encoding distinctive information depend on which of these two mechanisms are deemed responsible for false-memory reduction. In the following paragraphs, each mechanism is described along with the corresponding prediction for children’s performance.

Use of Item-Distinctive Information

Smith and Hunt (1998) argued that successful discrimination between studied and unstudied items at retrieval relies heavily on differential processing (i.e., distinctive processing) at encoding. Accordingly, any mental operation at encoding that assigns something distinctive to each DRM list member should lead to improved memory for studied items by supporting item-specific recollections and reduced false memory. Item-distinctive information is also believed to prevent relational processing (i.e., processing of similarities among items resulting in increased false memory). To test their hypothesis, Smith and Hunt (1998, Experiment 3) instructed experimental-condition participants to assign pleasantness ratings to each DRM-list word under the assumption that these ratings would allocate a distinctive characteristic (i.e., a different rating) to each word. Consistent with the authors’ prediction, participants who made pleasantness ratings showed significantly less false recall of critical lures than did participants who were merely asked to remember the words.

Effects of distinctive processing on children’s memory have been investigated, albeit not with the DRM task. For example, Howe et al. (2000) examined effects of distinctiveness manipulations on 5- and 7-year-olds’ memory. In one experiment, children studied word-pairs of object names. Each pair was introduced with a description of how the objects would interact. The interaction was either common or bizarre. For example, with the word-pair *dog–bicycle*, the experimenter verbally described either the common interaction “the dog chases the bicycle” or the bizarre interaction “the dog rides the bicycle.” Howe et al. found that for both 5- and 7-year-olds, memory performance was higher in the bizarre condition than in the common condition.

Other studies confirm that even young children’s memory can benefit from distinctive processing. Such benefits are apparent in studies of children’s memory for bizarre imagery compared with common imagery (Emmerich & Ackerman, 1979), script-violating information compared with script-consistent information (Hudson, 1988; but see Farrar & Goodman, 1992), and personally performed events compared with observed events (Baker-Ward, Hess, & Flannagan, 1990). If false-memory reduction is mainly based on enhanced discrimination between studied and unstudied items because of distinctive encoding of studied items, even young children’s false memory may decrease as a result of encoding distinctive information.

Use of the Distinctiveness Heuristic

Another explanation for the false-memory decrease due to distinctiveness was proposed by Schacter et al. (1999) after they observed that robust false-recognition effects were reduced substantially when adults studied each member of the DRM lists accompanied by a picture representing its meaning (Israel & Schacter, 1997; Schacter et al., 1999). In contrast to Smith and Hunt’s (1998) hypothesis, Schacter et al. (1999) proposed that the
suppression effects were not attributable to an enhanced memory-based discrimination ability but to a mental operation involving the use of a distinctiveness heuristic, a mode of responding based on participants’ metamemorial awareness that true recognition of studied items should include recollection of distinctive details: Because participants know that they have been provided with distinctive information for each word (i.e., pictorial stimuli), they demand access to recollection of distinctive information to endorse an item, thus becoming less likely to endorse an item unless they can retrieve “having seen the picture.” According to Schacter et al., this mode of responding is a strategic decision rule resulting in a conservative response bias.

Children’s decision processes in memory tasks have been infrequently studied (but see Berch, 1977; Brady, Poole, Warren, & Jones, 1999; McBrien & Dagenbach, 1998). Of particular relevance here is that a paucity of research exists on children’s ability to strategically adopt a conservative response bias based on the metamemorial awareness of the memory task demands. Although it has been shown that kindergartners may appropriately shift their response bias in a recognition memory test as a result of trial-by-trial feedback, they do not seem to succeed in changing their response bias as a result of general knowledge about how to maximize their performance in a memory task (Berch, 1977). Thus, the young children involved in the present study (5-year-olds) would not be expected to use a decision-based process such as the distinctiveness heuristic. If the distinctiveness heuristic is mainly responsible for reduced false recognition following encoding of distinctive information, an Age × Encoding Condition (i.e., distinctive vs. nondistinctive encoding) interaction should emerge such that young children should be equally likely to falsely recognize the critical lures when encoding distinctive information and when studying words alone, whereas older children and adults should be less likely to falsely recognize critical lures after distinctive encoding than after nondistinctive encoding.

The Subjective Experience of True and False Memories in Children and Adults

Introspection about their own mental states has been utilized to differentiate and characterize the processes involved in adults’ memory performance. For instance, one line of evidence that supports the dual-process (familiarity and recollection) account of recognition memory derives from studies in which remember–know judgments were elicited from participants (e.g., Gardiner & Java, 1993; Tulving, 1985; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998). In these studies, when participants recognized an item as old, they were asked to report (a) whether they remembered seeing the item (or hearing it, according to the presentation modality), that is, whether they recollected specific features of the item and/or of the study episode, or (b) whether in the absence of specific recollection they felt that they knew the item had been studied. The remember–know (R-K) procedure has also been used in false-memory studies (Hyman, Gilstrap, Decker, & Wilkinson, 1998; Roediger & McDermott, 1995) to assess whether false memories are phenomenologically similar to true memories. For example, Roediger and McDermott (1995) found that adult participants who falsely recognized the critical lures as old were more likely to label their false-memory experience as “remember” than as “know.” Many experiments that employed the R-K procedure for non-DRM lists have shown that false alarms are most often judged as “know” experiences (e.g., Jones & Roediger, 1995), suggesting that such false alarms were experienced as a sense of familiarity for the item in the absence of conscious recollection of specific details of the item. This is in quite sharp contrast to previous findings that false memories of critical lures in the DRM task tend to be described by adults as reflecting conscious recollection.

In contrast to the often-made attempt to evaluate phenomenological memory experience in adults, it has not been well established whether it is even possible to investigate phenomenological characteristics of memory in children. Some researchers have claimed that until children are a certain age, they may not differentiate between belief, knowledge, and memory, making it impossible to experience different memory states such as remembering and knowing. For example, Perner and Ruffman (1995) argued that young children cannot experience actual remembering of events, as differentiated from knowing that events happened to them, until they fully appreciate that what they know originated from previous personal experience. From a series of experiments on children’s understanding of the perceptual origin of their own knowledge (e.g., Perner & Ruffman, 1995; Wimmer, Hogrefe, & Perner, 1988; Wimmer, Hogrefe, & Sodian, 1988), the authors concluded that autonoetic representation may not be available to children until they are 3–6 years old, because up until that age, children do not appreciate the connection between knowledge and its access. Along this line, Lyon and Flavell (1994) found that 4-year-olds but not 3-year-olds understood the necessity of prior knowledge in order to remember and forget. Taken together, these studies suggest that children’s ability to experience actual remembering may develop between the ages of 3 and 5 years.

Although these studies provide valuable information as to the emergence of the experience of remembering in children, they do not clarify whether and when children can introspect on and make direct evaluations of their own memory experiences. Such clarification may be complicated by the fact that even if children possess some ability to introspect on their own mental states, they may still have difficulties in understanding and following verbal instructions such as the ones used to elicit R-K judgments. Thus, a method like the R-K procedure may not be suitable for studying young children’s phenomenological experience of memory. Alternative ways are therefore needed to explore qualitative characteristics of false memories in children. As explained next, source attributions and confidence ratings for true and false memories may provide valuable information for investigating children’s memory experiences.

Source Attributions

Source judgments afford a reasonable alternative for examining whether children actually remember as opposed to merely know that an event occurred. In R-K experiments with adults, participants are instructed to choose “remember” only when they can mentally re-experience the word presentation and are able to retrieve specific, qualitative information related to the presentation of the word (for example, a physical characteristic associated with the presentation of the word, what the word made them think of, or
what they were doing when the word was presented). Information about source constitutes a piece of qualitative information. Thus, although the range of qualitative information retrieved about a remembered item may transcend source, the source attribute to an item can be considered one valuable indicator of a recollective experience.

Analysis of source attributions of critical lures has already been used with adults to explore whether true and false memories are experienced differently (Lampinen, Neuschatz, & Payne, 1999; Mather et al., 1997; Payne et al., 1996). Some researchers found a few subtle differences in source attributions between studied items and critical lures, such as participants being slightly less willing to attribute a source to the critical lure than to studied items (87% vs. 94%; Payne et al., 1996) or being more willing to change their minds about the source assigned to the critical lure than about the source assigned to the studied items (Lampinen et al., 1999), which suggests that true and false memories may be experienced differently.

However, Mather et al. (1997) found that when all the words in a list were spoken by the same speaker and there were two speakers alternating across lists, participants were as likely to attribute the correct source to studied words as to the critical lure. For studied words, the correct source was the voice that actually spoke that word, whereas for critical lures, the “correct” source was the voice that spoke the words in the list converging on that critical lure. The authors concluded that false memories for lures were affected by the characteristics of the memories for the associated studied items. Specifically, Mather et al. proposed that the source attribution to the critical lure resulted either (a) from associating with the activated lure the characteristics of the studied items during encoding or (b) from inference processes at retrieval (i.e., participants would infer that they should attribute a certain source to a critical lure because they had assigned that source to the lure’s associates). In the present study, we used Mather et al.’s source task to investigate the quality of false memories across age groups. Given that children’s ability to correctly attribute a source to experienced events increases with age (e.g., Foley & Johnson, 1985; Foley et al., 1983), if endorsed critical lures and studied items are remembered similarly, then the rate of “correct” source attribution should vary according to age. Furthermore, to the extent that true and false memories are distinguishable for both children and adults, there should be a significant difference in the rates of “correct” source attribution between studied items and critical lures.

Confidence Ratings

Another way to characterize the quality of memory experiences is through eliciting confidence ratings. Although confidence ratings are influenced by factors extraneous to memory functioning per se (e.g., others’ positive feedback does not affect memory processing but tends to elevate witness confidence; Wells & Bradfield, 1999) and thus cannot be considered pure measures of memorial experience, confidence ratings are also believed to reflect memory processes. For example, such ratings are often obtained in recognition memory research to investigate the contribution to performance of two different processes, recollection and familiarity (Yonelinas, 1997, 2001). In false-memory research, Roediger and McDermott (1995) found that participants’ confidence was high after endorsing critical lures. This result was interpreted as an indication of the vividness of false memories elicited with this paradigm. To date, there is a paucity of research examining the relation between confidence and accuracy in children. Berch and Evans’s (1973) study of kindergartners’ and third graders’ performance in a continuous recognition memory task is a unique exception. For both age groups, it was found that the lower the child’s level of confidence in judging an item as old, the lower was the probability of that item’s actually being old. However, third graders were better than kindergartners at gauging the accuracy of their recognition responses; that is, compared with kindergartners, third graders were more confident when they recognized a studied item and less confident when they falsely recognized a nonstudied item. Nevertheless, even children as young as 5 years were found to be somewhat capable of introspection about their own memory states.

Thus, we exploited this ability to determine whether confidence ratings when endorsing critical lures would differ from confidence ratings when endorsing studied items. If, across age groups, false recognition for critical lures is similar to true recognition for studied items, confidence ratings should not vary across endorsed studied items and critical lures. If individuals experience true and false memories differently, confidence ratings assigned to true memories should be higher than confidence ratings assigned to false memories.

The Present Study

In the present study, we investigated age trends in spontaneous false memories with the DRM paradigm by testing 5- and 7-year-old children and college undergraduates. These children’s ages were chosen because they mark the beginning and the end of the “five to seven year shift” (for reviews, see Sameroff & Haith, 1996). The many transitions occurring during this period (e.g., from perceptual to conceptual classifications) may be relevant for the development of the false-memory effect as elicited with the DRM paradigm.

In addition, the effects of processing distinctive information in reducing false-memory formation were studied across age groups. For each age group, half of the participants studied words along with pictures, whereas the other half studied words alone. Finally, the phenomenological characteristics of true and false memories were subjected to examination. Data on source attributions and confidence ratings were gathered for true and false memories.

Method

Participants

Participants included 48 children from two age groups (5-year-olds, mean age = 64.79 months, range = 61–70 months; 7-year-olds, mean age = 88.29 months, range = 82–92 months) and 24 adults (mean age = 21.38 years, range = 18–37 years). Participants were randomly assigned to each experimental condition with the only constraint being that each age by condition cell include equal numbers of male and female participants. Approximately 79% of the participants were European American, 12% were Asian American, 5% were Hispanic, 3% were African American, and 3% were Native American. Participants were mostly from middle-class
backgrounds and were native English speakers. Children were recruited by advertising the research project in a local newspaper, whereas adults were recruited from introductory psychology classes. Child participants were compensated with a small prize and $5, whereas undergraduate students received class credit.

Materials

Word lists. Fifteen lists were selected from the materials used by Israel and Schacter (1997). Lists were selected on the basis of new adult norms for DRM lists (Stadler, Roediger, & McDermott, 1999) so that they produced variable intrusion rates of the critical lure. Every list included seven associate words (instead of the more typical 12 associate words) because of the limited immediate word span of young children (Dempster, 1981). The words selected were the strongest associates of the critical lures that fell within the vocabulary of 5-year-olds (Gilhooly & Logie, 1980). Use of seven associates per list guarantees reliable false recall and false recognition in adults, although the magnitude of the effects is somewhat reduced (Robinson & Roediger, 1997). The 15 lists were divided into three blocks of five lists each. Each participant studied words from two blocks (i.e., 10 lists), whereas the words from the third block were used as distractors in the recognition test. Target and distractor words and block order were counterbalanced within each age group and experimental condition.

The recognition test included 60 items: 30 studied and 30 unstudied. We created the recognition memory test similarly to the way in which Israel and Schacter (1997) created their memory tests. The 30 studied items, or true targets, were obtained by selecting three items from each presented list (always those in Serial Positions 1, 3, and 7). Three types of nonstudied items were also presented: (a) 10 critical lures of studied lists, or false targets; (b) 15 items from the five nonstudied word lists, or true target controls (these items were obtained by selecting three items from each nonstudied list, always those in Serial Positions 1, 3, and 7); and (c) 5 critical lures from the five nonstudied lists, or false target controls. The recognition test did not vary according to whether studied words and critical lures were reported during the recall phase.

Pictorial stimuli (Israel & Schacter, 1997). These stimuli were black- and-white line drawings of list items. Each drawing contained approximately equivalent amounts of detail.

Confidence rating board (adapted from Berch & Evans, 1973). This board was used to help child participants express their confidence. Two photographs, either representing a girl if the participant was female or a boy if the participant was male, were available. One picture depicted a child with a very confident facial expression, and the other depicted a child with a very doubtful facial expression. The two pictures were positioned at opposite ends of a board. Three dots were placed in the space between the two pictures, in reference to a 3-point scale (not confident, somewhat confident, and very confident).

Procedure

Parents of child participants were fully informed of the experimental procedures and signed the consent form upon arrival at the university laboratory. Adult participants also signed a consent form, which introduced the study as a list-learning task. The study was presented in the same way to child participants but in a developmentally appropriate manner.

Participants were tested individually. Overall, the experimental session consisted of two phases. In the first phase, participants were exposed to the word lists presented one at a time and attempted to recall the words after having heard each list. During the second phase, participants engaged in a recognition memory test.

There were two encoding conditions, namely, no picture and picture. In the no-picture condition, participants were told by the experimenter that they were taking part in a memory experiment and that they should try to remember the material as well as possible. Participants were also told that to perform well, they were to listen very carefully to the words. In the picture condition, participants were additionally told that to perform well, they had to look at the pictures.

Then participants watched a videotape in which two research assistants, Roger and Debbie, read five lists each (i.e., one block). The order of the male and female readers was counterbalanced. In the no-picture condition, the research assistant was filmed while simply saying a word every 5 s, whereas in the picture condition, the research assistant would also show a picture of the word while saying it (again, every 5 s). After each list, participants were involved in a filler task for about 25 s. To prevent rehearsal, during this interval, we requested 5- and 7-year-olds to say what number came before or after the number that the experimenter said, whereas we required adults to count backward by 7 starting from different numbers. At the end of the interval, participants were asked to recall all the words they remembered from the list without guessing.

After the presentation of the 10th list and recall test, participants were told that the experimenter would read words and that the participant would have to say “yes” if the word was one of those previously heard on the videotape and “no” if the word was not one previously heard. The administration of the yes–no recognition test was preceded by a careful explanation on how to use the 3-point confidence rating board. Specifically, participants were told that the child represented in the pictures was playing the same memory game they were about to play. Then participants’ attention was directed to the picture representing a confident face, and participants were told that sometimes that girl (or boy according to the participant’s gender) was really sure of having heard a word (or really sure of not having heard a word). Participants were instructed to point to the dot right next to the confident face if they felt sure, that is, so positive that they did not even have to think about their answer (Berch & Evans, 1973). Next, participants’ attention was directed to the picture representing the uncertain face, and children were told that sometimes the pictured girl or boy was not sure at all of having heard a word (or not sure at all of not having heard a word). Then participants were instructed to point to the dot right next to the uncertain face if they felt very unsure, that is, if they had to think really hard about the answer and felt that they were almost guessing. Finally, participants’ attention was directed to the middle dot, and they were told that they were to select the middle dot if they had to think a little before answering and if they did not feel as positive as they would when selecting the high-confidence dot but did not feel as confused as they would when choosing the low-confidence dot.

Participants’ understanding of these instructions was then tested. First, participants were asked questions such as “Which dot would you pick if you were really sure that you heard a word?” or “Which dot would you pick if you were not at all sure that you heard a word, and you thought really hard before saying that you heard it?” Then participants were shown pictures of animals or objects, one at a time. The pictures were either clear, partially covered with spots, or almost totally covered with spots. The levels of clarity were designed to elicit different levels of confidence (i.e., high confidence, middle confidence, and low confidence) for the identification of the object. One example was provided for each level of cleanness. Participants were asked yes–no questions about what was represented in the pictures (e.g., “Is this a plane?”). The correct response for half of the questions was “yes,” and for the other half it was “no.” After each response, participants were also asked to indicate their confidence using the confidence rating board. If participants made mistakes, the experimenter gave feedback and explained the use of the board again. The recognition test did not start until participants showed understanding of the use of the confidence rating board. Although some 5-year-olds evinced some difficulty at first in understanding the use of the board, they understood correct board use after being provided with feedback and a few additional trials.
Next, the recognition memory test started. After responding, children and adults were asked to rate their confidence with the confidence rating board. After rating their confidence but only in the case of a “yes” response, participants were asked to report whether Roger or Debbie had said the word. Participants were provided with pictures of the research assistants. To obtain an objective record of the participants’ responses, we videotaped the experimental sessions.

Finally, children were thanked and given a prize. Although our procedure typically lasted between 1 and 1.3 hr, and thus was somewhat protracted for young children, all participants completed the experimental session.

Results

Preliminary analyses revealed that the results did not significantly differ according to gender, block, or block order. Thus, the data were collapsed across these variables.

A series of 3 (age group: 5-year-olds, 7-year-olds, and adults) × 2 (encoding condition: picture vs. no picture) × 2 (item type: studied items vs. critical lures) mixed analyses of variance (ANOVAs) was conducted on dependent measures derived from recall, recognition, source attribution, and confidence rating performance. All statistically significant effects are reported.

Free Recall

Table 1 presents the proportion of correctly recalled items (i.e., the number of correctly recalled words divided by the total number of studied words) and the proportion of falsely recalled critical lures (i.e., the number of recalled critical lures divided by the total number of lists), which were subjected to statistical analysis. The results revealed a significant main effect of item type, \( F(2, 66) = 380.88, p < .01, \eta^2 = .85 \), such that participants were more likely to recall studied items (\( M = .66 \)) than critical lures (\( M = .20 \)). Item type also significantly interacted with encoding condition, \( F(2, 66) = 4.91, p < .05, \eta^2 = .07 \), such that studying words accompanied with pictures did not significantly affect recall for studied items, \( F(2, 66) = 0.82, p = .37, \eta^2 = .01 \), but significantly affected recall of critical lures, \( F(1, 66) = 4.88, p < .05, \eta^2 = .07 \), such that participants who saw the pictures during the study phase were less likely to falsely recall the critical lures (\( M = .15 \)) than were those who did not see the pictures (\( M = .24 \)). If one observes the proportions (see Table 1), however, it is evident that the advantage of encoding condition is driven by the child participants’ false recall. On average, adults falsely recalled the same number of critical lures in both encoding conditions (i.e., .15).

A significant main effect of age also emerged, \( F(2, 66) = 23.63, p < .01, \eta^2 = .42 \). However, this effect was qualified by an interaction between age and item type, \( F(2, 66) = 45.30, p < .01, \eta^2 = .58 \), such that the age effect held only for recall of studied items, \( F(2, 66) = 117.36, p < .01, \eta^2 = .63 \), but not for recall of critical lures, \( F(2, 66) = 2.37, p < .10, \eta^2 = .07 \) (simple effects analysis). Regarding recall for studied items, Bonferroni planned comparisons revealed that all age groups significantly differed \( (p < .05) \) from each other: 5-year-olds, \( M = .43 \); 7-year-olds, \( M = .66 \); undergraduates, \( M = .89 \).

Because of the age effect in correct recall, a second measure of false recall was established, specifically, the number of critical lures recalled divided by the number of all words recalled (i.e., studied or nonstudied). When this proportion for each participant was used as a dependent measure in a 3 (age group: 5-year-olds, 7-year-olds, and adults) × 2 (encoding condition: picture vs. no picture) ANOVA, a significant age effect emerged, \( F(2, 66) = 11.78, p < .01, \eta^2 = .26 \). Bonferroni planned comparisons showed that 5-year-olds (\( M = .08 \)) were more likely than 7-year-olds (\( M = .04 \)) and adults (\( M = .02 \)) to falsely recall the critical lures. The effect of encoding condition was again statistically significant, \( F(1, 66) = 5.11, p < .05, \eta^2 = .07 \), such that participants in the picture condition were less likely to falsely recall the critical lures (\( M = .04 \)) than were participants in the no-picture condition (\( M = .06 \)).

Recognition

Table 2 shows the proportions of “yes” responses to true targets (i.e., studied items), true target controls (i.e., words belonging to nonstudied lists), false targets (i.e., critical lures of studied lists), and false target controls (i.e., critical lures of nonstudied lists) as a function of age and encoding condition. These proportions were used to derive corrected true and false recognition scores.

Consistent with the logic articulated by Israel and Schacter (1997), we obtained corrected true recognition scores by subtracting the proportion of “yes” responses to true target controls from the proportion of “yes” responses to true targets, to ensure that the observed true recognition rates were not a result of a general tendency to endorse items regardless of whether they were studied or nonstudied. Corrected false recognition scores were computed by subtracting the proportion of “yes” responses to false target controls from the proportion of “yes” responses to false targets, to ensure that the observed false recognition of critical lures actually resulted from studying the members of the lists and not from a general tendency to endorse critical lures (e.g., critical lures may generally evoke a high sense of familiarity) regardless of whether the members of the list were actually studied. Corrected true and

### Table 1

**Proportions of Recalled Studied Items and Critical Lures**

<table>
<thead>
<tr>
<th>Item type</th>
<th>5-year-olds</th>
<th>7-year-olds</th>
<th>Undergraduates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No picture</td>
<td>Picture</td>
<td>No picture</td>
</tr>
<tr>
<td>Studied items</td>
<td>.46 (.13)</td>
<td>.41 (.14)</td>
<td>.61 (.11)</td>
</tr>
<tr>
<td>Critical lures</td>
<td>.32 (.17)</td>
<td>.18 (.19)</td>
<td>.25 (.16)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses.
false recognition scores for each participant were entered in a 3 (age group: 5-year-olds, 7-year-olds, and adults) × 2 (encoding condition: picture vs. no picture) × 2 (item type: studied items vs. critical lures) mixed ANOVA as described above.

There was a significant main effect of item type, \( F(1, 66) = 324.22, p < .01, \eta^2 = .83 \), such that true recognition (\( M = .80 \)) was significantly higher than false recognition (\( M = .29 \)). The effect of item type significantly interacted with encoding condition, \( F(1, 66) = 12.44, p < .01, \eta^2 = .16 \), such that picture encoding enhanced corrected true recognition scores (picture, \( M = .84, SD = .12 \); no picture, \( M = .76, SD = .15 \)), \( F(2, 66) = 8.75, p < .01, \eta^2 = .12 \), and decreased the false-recognition effect (picture, \( M = .23, SD = .17 \); no picture, \( M = .35, SD = .25 \)), \( F(1, 66) = 5.47, p < .05, \eta^2 = .08 \).

Finally, the interaction between item type and age approached statistical significance, \( F(2, 66) = 2.87, p = .06, \eta^2 = .08 \). Justified by our predictions, simple effects analyses were conducted. These analyses revealed that for true recognition scores, a significant age effect was found, \( F(2, 66) = 12.25, p < .01, \eta^2 = .27 \). Bonferroni planned comparisons revealed that adults had significantly higher true recognition scores (\( M = .90, SD = .08 \)) than did 5-year-olds (\( M = .73, SD = .15 \)) or 7-year-olds (\( M = .78, SD = .14 \)). The difference between 5- and 7-year-olds’ true recognition scores, however, was not significant. For false recognition scores, however, there was no significant effect of age, \( F(2, 66) = 0.62, p = .54, \eta^2 = .02 \).

In summary, participants were more likely to form true memories than false memories. In addition, although true memory performance not surprisingly increased with age, false-memory formation was generally not affected by age (except for one measure of false recall). Finally, studying pictures improved memory performance; that is, it increased true recognition and decreased false recall and recognition.

### Source Attributions

Each participant’s proportion correct source-attribution score was entered in a full-model 3 (age group: 5-year-olds, 7-year-olds, and adults) × 2 (encoding condition: no picture vs. picture) × 2 (item type: studied items vs. critical lures) mixed ANOVA. A significant age main effect, \( F(2, 57) = 8.34, p < .01, \eta^2 = .23 \), was qualified by a significant Age × Encoding Condition interaction, \( F(2, 57) = 3.93, p < .05, \eta^2 = .12 \) (see Figure 1). Simple effects analyses revealed that 5-year-olds were better at attributing the correct source when they studied words without pictures rather than with pictures, \( F(1, 21) = 11.17, p < .01, \eta^2 = .35 \), whereas the other age groups’ source-attribution performance did not significantly differ according to encoding condition, \( F(1, 21) < 0.16, ps > .70 \). Also, the simple effect of age on source attributions was significant only for participants in the picture condition, \( F(1, 28) = 10.30, p < .01, \eta^2 = .32 \), and not for those in the no-picture condition, \( F(1, 29) = 1.03, p = .37, \eta^2 = .02 \). Bonferroni planned comparisons for source-attribution performance in the picture condition showed that 5-year-olds were significantly less accurate than 7-year-olds and undergraduates. The performances of the last two groups did not significantly differ.

No significant effect of item type was found, \( F(1, 57) = 0.65, p = .42 \). Thus, source-attribution accuracy did not vary according to whether participants were attributing a source to studied items (\( M = .72, SD = .14 \)) or critical lures (\( M = .69, SD = .28 \)), reflecting similar phenomenological experience for true and false memories.

### Confidence Ratings

Analysis of confidence ratings assigned by participants when recognizing studied words and critical lures as “old” provides an additional indication as to whether participants had similar phenomenological experiences for both types of items. For each participant, mean confidence ratings for endorsed studied items (\( M = 7.26, SD = 1.32 \)) and critical lures (\( M = 7.18, SD = 0.27 \)) were calculated. Table 3 presents the mean confidence ratings. A significant main effect of item type emerged, \( F(1, 57) = 21.32, p < .01, \eta^2 = .27 \), such that studied items (\( M = 1.78, SD = .27 \)) were endorsed more confidently than critical lures (\( M = 1.54, SD = .42 \)). As is evident from Table 3, confidence for endorsed studied items and for critical lures did not differ for 5-year-olds in the no-picture condition. In addition, a significant main effect of encoding condition revealed that participants endorsed words more confidently when they studied words accompanied by pictures (\( M = 1.74, SD = 0.27 \)) than when pictures were not studied (\( M = 1.59, SD = 0.39 \)), \( F(1, 57) = 4.86, p < .01, \eta^2 = .08 \).

\(^1\) When \( A’ \) values for true and false recognition were subjected to statistical analysis, results replicated the pattern found with corrected recognition scores.
Three main goals motivated the present study: The first was to examine developmental patterns in the rate of false memories produced with the DRM paradigm; the second was to study whether false memories are reduced when children process distinctive information; and the third was to investigate children’s and adults’ subjective characteristics of true and false memories.

Developmental Differences in Spontaneous False-Memory Formation

Extant literature has not provided a firm basis for clear-cut predictions about developmental patterns of false-memory formation on the DRM task. On the basis of current theory, both age-related increases and decreases in false memories were plausible expectations. Our results showed no significant age differences in false-memory formation in recall or in recognition, suggesting that under the conditions examined in the present study, children do not differ from adults in the degree to which they create false memories for critical lures. However, adults outperformed children in the number of items correctly recalled and recognized. When this age effect was taken into consideration, younger children were more prone to false recall of critical lures. Because the probability of creating a false memory with the DRM paradigm increases in relation to the strength of association among the studied items, this result is consistent with the idea that, at least with the material used in the present study (i.e., highly frequent and concrete words), there were no obvious age-related differences in levels of associations between the studied words and the critical lures. Such differences may have been implicated had we found an increase of the false-memory effect with age.

Our data seem consistent with the general notion that as the abilities to recall and recognize previously encountered material and to extrapolate meaning from such material develop, so do the abilities to monitor the quality of memories and to discriminate between them and other memory traces, even those that share considerable similarities with studied items and that are strongly associatively activated across age groups (e.g., Johnson et al., 1993).

Since our research was conducted, other developmental studies with the DRM paradigm have been reported (Brainerd, Reyna, & Forrest, in press; Price, Metzger, Williams, Phelps, & Phelps, 2001). Brainerd et al. (in press) conducted a series of experiments in which 5-, 7-, and 11-year-olds and adults studied DRM word lists of 12 (or 15) items each. These researchers found that older children and adults showed much higher rates of false memory, particularly false recall, than younger children. Brainerd et al. interpreted their data as demonstrating that 5-year-olds failed to extract the gist of DRM materials. However, in our study, young children demonstrated the same false-memory level as (and in false recall, a proportionally higher level than) older children and adults. It should be added that when we compared performance across the two studies, adults’ false recall level was higher in the Brainerd et al. (in press) research than in our study (.37 vs. .15). The same was true for false recognition (.65 vs. .32). Why are Brainerd et al.’s results different from ours? One difference between our experimental procedure and the one used by Brainerd et al. is in the number of associate words included in each list. It is known that list length is reliably associated with the DRM false-memory effect such that as the number of associates increases, so does the probability that the critical lure will be falsely remembered (Deese, 1959; Robinson & Roediger, 1997). According to FTT, lists including a high number of associates are more likely to elicit the formation of gist traces than are those including only a few associates. Because we included 7 associates, whereas Brainerd et al. included 12 or 15, FTT’s prediction is consistent with the difference in adults’ false memory observed between our study and that of Brainerd et al.: Brainerd et al.’s lists may have resulted in higher gist extraction, thus leading to higher false-memory levels.
This account does not, however, explain why 5-year-olds in our study produced a higher level of false recall than did their counterparts in Brainerd et al.’s (in press) study (.25 vs. .06). One question arising from this comparison is whether the mechanisms underlying the false-memory effect for children are the same as those for adults. Robinson and Roediger (1997) cleverly demonstrated that contrary to Deese’s (1959) proposition, the false-memory phenomenon is predicted by the sum of the associative strengths of the studied items instead of the mean associative strength. This is an important finding because it implies that if an item that is only weakly related to the critical lure is added to the word list at study, it will still increase the sum of the associative strengths (whereas, being weakly associated, it reduces the mean associative strength) and the likelihood that the critical lure will be recalled. As noted above, 5-year-olds in our experiment produced higher levels of false recall after studying 7 items than did 5-year-olds in Brainerd et al.’s experiment after studying 12 (or 15) items. It is possible that as the number of associates increases, either because researchers include items that are only weakly associated with the critical lure or because they include lower frequency words, activation becomes more “diluted” than potentiated in children. This possibility is consistent with the findings of the other DRM study conducted with child participants (Price et al., 2001). Price et al. presented 7-year-olds and undergraduate students with 8-item DRM lists. In contrast to the usual DRM procedure, thematic lists were not presented one at a time, but two at a time during the study phase. Thus, participants were actually presented with lists of 16 words each. As in Brainerd et al.’s study, children obtained lower levels of false memory than did adults. Interestingly enough, adult participants in Price et al.’s study showed a false-recall level comparable to that found in our study (.19 vs. .15), which is reasonable given that participants in Price et al.’s experiment studied a number of associates per list that was comparable to that studied by our participants (8 vs. 7). However, 7-year-olds in our study showed higher false recall than 7-year-olds in Price et al.’s research (.19 vs .02). It is possible that studying together two 8-item lists of different themes “diluted” the activation of the critical lures for children, leaving unaltered the activation in adults. Future studies should further investigate the nature of developmental differences in the false-memory effect.

Finally, it should be added that studying lists of 12 or more items may be too demanding for young children, as was demonstrated in Brainerd et al.’s (in press) study by 5-year-olds’ low levels of correct recall, and quite low levels of recognition as well, in contrast to older participants’ much higher rates of true and false memories. In Brainerd et al.’s study, even assuming that the level of activation was the same across age groups (which may not be a valid assumption given that, compared with shorter lists, 12-item lists include weaker associates and lower frequency words that may produce different activation across age groups), it would still be reasonable to expect developmental differences in correct and false recall that were due to differential retrieval abilities across age groups and were exacerbated by the amount of to-be-retrieved material. Because developmental differences in memory ability (e.g., retrieval abilities) may have differentially contributed to false-memory formation across age groups, the developmental increment in the ability to extract gist may have been overestimated in Brainerd et al.’s study.

Nevertheless, in line with Brainerd et al.’s (in press) proposal, it is possible to design experimental conditions in which older children and adults are more likely than younger children to falsely remember critical lures, such as conditions in which access to verbatim traces may be prevented or reduced across age groups and the retrieval of gist traces encouraged. In this case, older children’s and adults’ further-developed abilities to connect each item to a theme could be exploited to promote higher false-memory levels for them.

On the other hand, adults’ ability to discriminate between true and false memories may have been enhanced by some characteristics of our experiment. In our study, the overall level of false memory was smaller than in other studies. As Robinson and Roediger (1997) demonstrated, the likelihood that individuals will falsely recall and recognize critical lures is correlated with the sum of the associative strengths of the list members. It follows that the more list members that are studied, the more likely it is that false memory will be experienced. Because in the present experiment participants studied seven items per list, the sum of the associative strengths was relatively limited, thus facilitating the discrimination between true and false memories. However, presenting children with seven words at a time was advantageous for maintaining young children’s cooperation throughout the experiment and guaranteed acceptable levels of performance across age groups. In addition, because the encoding manipulation in the present study required the words to be easily represented pictorially, the lists used in the present study included mainly concrete nouns. This has already been shown to reduce the false-memory effect compared with lists containing words that are not easily represented pictorially, even when larger numbers of items per list are used (Israel & Schacter, 1997). To the extent that participants spontaneously evoke mental images of the stimulus words and that concrete nouns facilitate this process, participants who study concrete nouns would have more distinctive information available later on (i.e., this is like a weaker version of the picture encoding condition).

Distinctive Information and False-Memory Reduction

Earlier we presented two accounts that have been proposed to explain the reduction of false memories after distinctive encoding. According to Smith and Hunt (1998), because distinctive encoding results in specific features being associated with each studied item, the discrimination of studied items and unstudied items is facilitated. In contrast, Schacter et al. (1999) proposed that the false-memory decrement results from implementing a heuristic that allows for rejection of critical lures when individuals are aware that they should demand retrieval of distinctive information to endorse an item.

In the present study, the presentation of pictures along with words reduced false-memory formation across age groups. In addition, participants were more likely to correctly recognize studied words if they were encoded with pictures. Given young children’s limited overt and spontaneous metacognitive and strategic skills, we consider it unlikely that 5-year-olds (and even 7-year-olds) implemented a conservative decision rule based on the explicit metamemorial awareness of characteristics of the task, as proposed by Schacter et al. (1999). The explicit and systematic nature of the distinctiveness heuristic is what makes its use un-
likely among young children. This is not to say that we uncovered evidence against the existence of a distinctiveness heuristic. Our results, however, suggest that event distinctiveness may reduce false-memory formation even in the absence of systematic strategy use.

What are other potential mechanisms for false-memory reduction after encoding pictures? Smith and Hunt (1998) proposed that the availability of distinctive information for encountered words facilitates discrimination between old and new items by enhancing item-specific processing (i.e., processing of the differences among items) and reducing relational encoding, the latter of which emphasizes the relations among list items and therefore possibly enhances false-memory formation. However, the mental processes through which item-specific information prevents false memory were not further specified.

FIT describes a possible mechanism for correct rejection of stimuli sharing high similarity with studied stimuli. For instance, Brainerd, Reyna, and Kneer (1995) found that for 5- and 8-year-olds, when retrieval of verbatim traces was supported through the presentation of studied items before the presentation of related distractors, false recognition decreased. In contrast, when related distractors were presented in the test list before the studied items, false recognition increased. This result was interpreted as evidence that correct rejection may be supported by mental processes called nonidentity judgments: After encountering a studied item at test and then having verbatim traces made available to them, individuals were better at recognizing that the distractors were only similar to, but not the same as, studied stimuli, thus enhancing discrimination between studied material and nonstudied material.

Recently, Rotello, Macmillan, and Van Tassel (2000) proposed, more generally, that recall mechanisms may be responsible not only for endorsement of studied stimuli but also for rejection of nonstudied stimuli. For instance, they proposed that a recall-to-reject process supports the rejection of distractors (e.g., the critical lure sleep in the DRM paradigm). According to Rotello et al., at test, a distractor may serve as a retrieval cue for similar items that were originally studied. Individuals then evaluate the recalled items against the distractor. If, on some set of features, the distractor does not match the originally studied items that are recalled, then the distractor is rejected. Thus, recalling studied items helps individuals reject the distractor.

In our experiment, when participants studied words with associated pictures, participants were provided with an additional feature on which to base the discrimination between the critical lure and the studied items. Thus, according to the recall-to-reject hypothesis, individuals in the picture condition would be better able to reject critical lures and, as a result, to diminish false memories. To the extent that the recall-to-reject process may explain our results, it implies that children as well as adults can utilize this mechanism. Interestingly enough, it is plausible that the recall-to-reject process could, consistent with Schacter et al.’s proposal, produce a conservative decision style of responding, as reflected in higher correct rejection rates. Such a decision process, however, may be conceived as a by-product of the recall-to-reject process instead of as the inherent expression of a systematic heuristic. Except for a few important exceptions (e.g., Brainerd et al., 1995), the developmental patterns of processes supporting correct rejection have rarely been studied. In fact, specific mechanisms involved in correct rejection have rarely been studied at large (but see, e.g., Brainerd, Reyna, & Mojardin, 1999; Rotello & Heit, 2000; Rotello et al., 2000; Strack & Bless, 1994). The literature on false-memory formation would benefit greatly from investigation of this important issue.

Further research should also explore whether there is a developmental change in how item-distinctive information is used generally. In the present study, we found that pictures representing the words helped participants’ accuracy. According to Smith and Hunt (1998), however, any type of information differentiating items from each other (e.g., degrees of pleasantness) would contribute to false-memory reduction. Whether this is true for children is of interest for an understanding of the potential limits of children’s use of distinctive information to enhance accuracy and reduce error. For instance, Howe et al. (2000) found that kindergartners involved in a Von Restorff paradigm did not show enhanced memory for a numerical isolate among a majority of objects (e.g., clothing). The authors interpreted their result as due to the fact that the isolate may have been too dissimilar from the other items to provide an advantage in children’s memory performance. It is conceivable that encoding pictures in the present experiment helped children because the information provided was consistent with the meaning of the words. Whether arbitrary relations between studied words and distinctive information encoded along with them would still promote false-memory reduction is an empirical question. Answering this question may provide valuable information for understanding, from a developmental perspective, the reasons why distinctive encoding enhances memory.

Subjective Experience of True and False Memories

One of the goals of the present study was to examine subjective experiences of true and false memories in spite of the methodological challenges imposed by working with young children. Comparing source attributions and confidence ratings for endorsed studied items and critical lures constitutes a first step in the investigation of children’s phenomenological experience of true and false memories.

We found that the accuracy of source attributions did not differ according to whether participants’ memory was true or false. Overall, source-attribute accuracy depended on age and on an unexpected interaction between age and encoding condition. Because there was nothing about the pictures that was source-specific, that is, the pictures did not vary systematically across the two sources, picture encoding was not expected to add anything to individuals’ ability to differentiate between the two sources. However, pictorial encoding hurt 5-year-olds’ performance. Thus, for 5-year-olds, paying attention to pictures (which later helped participants to discriminate between true and false memories) was at the expense of retaining other information, in this case external source-specifying information.

Failing to reveal a difference in source accuracy between studied stimuli and critical lures is interesting not only because it may suggest similar subjective experience for true and false memories across age groups but because it may provide information on the mechanisms supporting such attributions. As mentioned previously, Mather et al. (1997) argued that “correct” source attribution for endorsed critical lures could occur either because associations
with the lure come to mind during encoding or because of infer-
ence processes at test. This inferential process is quite complex:
When asked about the source for an endorsed critical lure, indi-
viduals remember the source of the studied items converging on
that critical lure and reason that if a certain person spoke the other
members of the word list, then that person must have spoken the
critical lure as well.

Could children as young as 5 years perform the type of inference
hypothesized by Mather et al. (1997)? Studies on young children’s
abilities to infer and use inferences as a source of knowledge show
that, under certain conditions, young children can perform logical
deduction (e.g., Dias & Harris, 1988; Pillow, Hill, Boyce, & Stein,
2000; Sodian & Wimmer, 1987). For example, in Sodian and
Wimmer’s (1987, Experiment 2) study, 4- and 5-year-olds were
shown two containers, each filled with one type of candy (only
yellow candies or only black candies). Children were then in-
formed that one piece of candy was taken from one of the con-
tainers and put in a bag, and they were told from which container
the candy was taken. Some 4-year-olds and the majority of 5-year-
olds demonstrated that they could correctly infer the type of candy
in the bag from knowing the container from which the candy was
taken.

In studies on inference processes such as the one just described,
children are introduced to verbal syllogisms whose premises are
generated by the experimenter and that refer to concrete objects
available for the child’s inspection. In contrast, an inference about
the source of a memory would involve premises that refer to
psychological states rather than concrete objects, and the premises
would be internally generated rather than generated by the
researcher.

Given the complexity of the inferential process hypothesized by
Mather et al. (1997), it is possible that inferring the source of a
memory might be relatively difficult for young children. If this is
the case, and because 5- and 7-year-olds showed comparable
source-attrition accuracy for true and false memories, it seems
most likely that source is assigned to the critical lure during the
encoding phase rather than as a result of inference processes.
Future research should be conducted to contrast the encoding and
inferential accounts of source attribution for false memories.

In contrast to source attributions, confidence ratings for en-
dorsed studied items were significantly different from confidence
ratings for critical lures; participants were generally more confi-
dent when endorsing studied items than when endorsing critical
lures. Nevertheless, confidence ratings of 5-year-olds in the no-
picture encoding condition did not significantly differ for studied
items and critical lures. Thus, these 5-year-olds did not experience
time and false memories differently, at least as indexed by confi-
dence ratings.

Encoding condition, however, similarly affected confidence for
both true and false memories: Participants’ endorsements for both
studied items and critical lures were expressed with higher confi-
dence when pictures accompanied the words than when words
were studied alone. That studying pictures increases confidence for
studied items may not be too surprising given that encoding of
visual information increased true recognition, possibly by making
memory for words more vivid. In contrast, it may be surprising
that studying pictures also increases confidence for endorsed crit-
cical lures: Because studying pictures reduces false recognition by

providing a better means to discriminate between true and false
memories, one might expect individuals in the picture condition
who still fall for a critical lure and falsely recognize it to do so less
confidently than individuals who did not see pictures at study. Our
results are not consistent with this expectation. Instead, our results
suggest that when pictorial information is not used to correctly
reject a critical lure, it may heighten the vividness, reflected in
high confidence, of the false memory. This heightened vividness
may result from an association at encoding between the activated
critical lure and the pictorial information accompanying the to-be-
remembered words.

Although encoding condition affected confidence for both true
and false memory, source attributions and confidence ratings in
general did not behave in a similar manner. In principle, these
tasks reflect different mental operations. When source judgments
were requested, participants had already recognized a word, and
they were required to distinguish between two external sources.
Confidence judgments were elicited from participants to assess
their certainty for the discrimination between internal and external
sources of information (i.e., Did I hear that word, or did I just think
about it?), the failure of which resulted in false recognition of
critical lures. Given that source-attrition and confidence tasks
tap two related but different aspects of the phenomenal experience
of false memory, it is not surprising that they did not behave
equivalently.

Conclusion

In the present study we investigated false-memory formation in
children and adults using an experimental paradigm, the DRM
task, that has formerly proven to be a successful tool for acquiring
deeper knowledge of false memory in adults and neurological
patients. We fruitfully used this paradigm with child participants
and determined that children as well as adults are likely to create
false memories for words representing the theme of a list and that
children, like adults, are less likely to form false memories if they
encounter distinctive information along with words than if they
study words alone. Equally important is that we have begun to
explore children’s subjective experience of true and false mem-
ories, and how children attribute qualities to their memories, by
requiring them to make source and confidence judgments. We
have therefore demonstrated that the DRM task is advantageous
for the study of several developmental questions pertaining to the
nature of false memories, the processes supporting rejection of
false events, and the subjective experience of true and false
memories.

References

involvement on children’s memory for events. Cognitive Development,
5, 55–69.
of the Psychonomic Society, 10, 328–330.

Received July 2, 2001
Revision received March 25, 2002
Accepted March 25, 2002

---

**Low Publication Prices for APA Members and Affiliates**

**Keeping you up-to-date.** All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

**Essential resources.** APA members and affiliates receive special rates for purchases of APA books, including the Publication Manual of the American Psychological Association, and on dozens of new topical books each year.

**Other benefits of membership.** Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

**More information.** Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.