This study investigated the development of all 3 components of episodic memory (EM), as defined by Tulving, namely, core factual content, spatial context, and temporal context. To this end, a novel, ecologically valid test was administered to 109 participants aged 4–16 years. Results showed that each EM component develops at a different rate. Ability to memorize factual content emerges early, whereas context retrieval abilities continue to improve until adolescence, due to persistent encoding difficulties (isolated by comparing results on free recall and recognition tasks). Exploration of links with other cognitive functions revealed that short-term feature-binding abilities contribute to all EM components, and executive functions to temporal and spatial context, although ability to memorize temporal context is predicted mainly by age.

According to its most recent definition, episodic memory (EM) refers to personal events recollected in the context of a particular time and place, and with some reference to oneself as a participant in the episode (Tulving, 2001). The typical development of EM has received surprisingly little attention despite its central role in structuring identity (Nelson & Fivush, 2004; Piolino, Desgranges, & Eustache, 2009) and its frequent impairment in common neurological disorders such as epilepsy (e.g., Helmstaedter & Elger, 2009).

Developmental Trajectories of the Different Components of EM

Although the earliest roots of EM can be traced back to around a child’s second birthday, major changes in EM skills occur between 2 and 6 years, with substantial development continuing thereafter until adolescence (Newcombe, Lloyd, & Ratliff, 2007). The age-related improvement in memorizing core factual information (e.g., objects, pictures) has been largely documented and observed whatever the nature of the task (free recall: McAuley, Brahmbhatt, & Barch, 2007; cued recall: Melinder, Endestad, & Magnussen, 2006; recognition: Chiu, Schmithorst, Brown, Holland, & Dunn, 2006).

By contrast, the ability to memorize the spatiotemporal context of core factual information has been largely neglected in the developmental literature. A few studies have nonetheless suggested that recalling the context and recalling the core factual information itself may reflect two processes that develop independently (Czernochowski, Mecklinger, Johansson, & Brinkmann, 2005; Naito, 2003; Romine & Reynolds, 2004). According to the dual-process model of memory (Yonelinas, 1999), the recall of core factual information would be supported by familiarity-based processes, which are relatively automatic and includes few details, and contextual information would be supported by
recollection-based processes, which allow the inclusion of many specific details of the event. Recent behavioral and electrophysiological studies have indicated that familiarity-based processes develop earlier than recollection-based ones, which undergo protracted development (Billingsley, Smith, & McAndrews, 2002; Friedman, de Chastelaine, Nessler, & Malcolm, 2010; Ghetti & Angelini, 2008).

Accordingly, a few studies have observed that the context may be much more complex for children than recalling the information itself (Cycowicz, Friedman, Snodgrass, & Duff, 2001). Even though spatial context and temporal context are regarded as the defining features of EM (Tulving, 2001), related memory abilities have been largely neglected in the developmental literature. Some researchers (e.g., Gulya et al., 2002) claim nonetheless that the ability to recall spatial information is particularly complicated for children to acquire and that performances continue to improve beyond 10 years. However, due to the nature of the methodologies that are generally used to assess recall of spatial context in the literature, a note of caution needs to be sounded about its putative developmental trajectory. The characteristics of the spatial contexts used in these studies are relatively far removed from real life, as they are not particularly distinctive (two squares in a grid are visually very similar), have no actual meaning, and are used redundantly (association of one context with several different core factual contents). The few studies have investigated the ability to recall temporal context also point to the involvement of a process that school-age children apparently find particularly complex (Friedman, 1993, 2005; McCormack & Hoerl, 1999).

In addition to good spatiotemporal contextual recall abilities, remembering a “true” EM appears to require other skills, namely, the long-term binding processes. If we are to take into account the cohesive nature of EM representations, it is necessary to assess the concurrent recall of core factual content and its contextual features, contrary to the many studies described above (e.g., Cycowicz et al., 2001). Adopting just such an approach, Sluzenski, Newcombe, and Kovacs (2006) observed an age-related improvement between 4 and 6 years of age in the ability to remember associations between core factual content (i.e., an animal) and spatial context (i.e., a landscape) but not in memory for the isolated parts. These findings suggest that long-term binding abilities are extremely age sensitive in preschoolers, but the time course of the development in school-age children and adolescents remains to be pinpointed within the overall developmental trajectory of EM. Furthermore, researchers have yet to assess the development of memory for spatial and temporal contexts, two features that are embedded in EM, within the same sample of children. It may be assumed that these two forms of associative memory (association of spatial or temporal context with core factual information) exhibit distinct developmental trajectories, especially in light of studies with adults pointing to separate encoding processes for each type of context (e.g., van Asselen, Van der Lubbe, & Postma, 2006).

**Underlying Mechanisms of EM Development**

The underlying mechanisms responsible for changes in EM abilities in young children are rarely examined and yielded mixed findings. There are two separate hypotheses: It may rely either on the gradual improvement in encoding abilities or, quite the opposite, on the development of retrieval processes. In support of the encoding deficit hypothesis, studies have shown that memory skills improve with age in tasks involving straightforward retrieval processes (short retention intervals, use of recognition tasks; Sluzenski, Newcombe, & Ottinger, 2004). Then again, Lloyd, Doydum, and Newcombe (2009) support the retrieval hypothesis as they observed age differences for combinations (pictures in various visual contexts) when “long” study test intervals were used rather than short ones. These mixed results indicate that more studies are required, especially as previous research has not paid attention to the potential differences arising from the nature of the EM component involved.

Further, the development of EM abilities may also stem from developmental changes in other cognitive functions. Indeed, there is a wealth of evidence supporting the idea that EM development may rely on the gradual improvement in short-term feature-binding abilities. Like long-term abilities, short-term feature binding is thought to emerge gradually up until adolescence (Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Lorsbach & Reimer, 2005). Moreover, investigations of short-term memory have also demonstrated that performances increase markedly during childhood when feature binding is required, whereas the effect of age is less important for single features (Lorsbach & Reimer, 2005; see Sluzenski et al., 2006, for similar finding in long-term feature-binding memory). Only one study has yet investigated memory for combined features with different lengths of time between encoding and retrieval (Lloyd et al., 2009). Unfortu-
nately, it was impossible to look directly for potential links between feature-binding abilities on the basis of the length of retention interval as each participant performed only one of the two tasks (short- or long-term memory test). Apart from short-term feature binding, improvements in executive functions (EFs) may also contribute to developmental changes in EM. This assumption is supported by the well-known relation between strategic processes and EFs, considering that EFs are known to improve continuously throughout childhood and adolescence (Huizinga, Dolan, & van der Molen, 2006; Kalkut, Han, Lansing, Holdnack, & Delis, 2009). In fact, researchers have demonstrated that performances on some EM abilities are positively related to EFs performance (e.g., memory for the color of various objects; Cycowicz et al., 2001). Nonetheless, the nature of the link between EFs and EM abilities in children has yet to be specified and compared with the potential role of short-term feature-binding abilities.

In sum, the above-mentioned studies confirm that childhood is a period of considerable development when it comes to memorizing core factual content. By contrast, the development of “true” EM, that is, the acquisition and recall of complex events situated in a specific spatial and temporal context, has yet to be the subject of an in-depth investigation. Context may be more sensitive to age than core content memory, but studies have focused on material associating information with only one context (very rarely on temporal context) and therefore of little ecological validity. Concerning the mechanisms behind these changes, the most recent studies have yielded mixed findings, supporting either the encoding or the retrieval hypothesis. The link with the development of other cognitive functions, namely, short-term feature-binding abilities and EFs, needs to be investigated, too. Our study was therefore the first to address the development of EM in its richest and strictest definition (based on Tulving’s, 2001, framework) and from an ecological viewpoint. A new test was designed to assess the ability to memorize specific, integrated events, close to real life, from childhood to adolescence (4- to 16-year-old participants). We wanted first to shed light on the developmental trajectories of each EM component and second to look at the mechanisms underpinning each of them. To examine the role of encoding and/or retrieval in EM development, we compared the effects of age on EM abilities in tasks involving several different retrieval processes (free recall, cued recall, and recognition). We also studied the link between the development of each EM component and the cognitive functions that are thought to contribute to it (i.e., short-term feature-binding processes and EFs).

Based on previous investigations, we hypothesized that age would have a considerable effect on EM abilities in free recall, notably for the recall of contextual information (e.g., Gulya et al., 2002). If this effect was due to the low efficiency of the retrieval process in younger children, subsequent cued recall and recognition tasks would not display any similar effect of age. Rather, making retrieval easier for younger children would allow them to attain EM performances comparable to those of older children. Encoding process deficits, on the other hand, would be attested by low scores, even on the recognition task (see Bowler, Gardiner, & Berthollier, 2004, for a similar reasoning). As even young children generally perform well on the recognition of factual information, retrieval difficulties were expected for core content, whereas additional encoding difficulties were expected for the two contexts (e.g., Friedman, 2005; Gulya et al., 2002). Given the similarity between the developmental trajectories of short- and long-term feature-binding abilities (see Lorsbach & Reimer, 2005; Sluzenski et al., 2006), we expected the development of EM abilities to be associated with short-term feature-binding abilities when different items had to be simultaneously recalled. Moreover, we hypothesized that the age-related improvement in EFs would be reflected in a gradual improvement in EM for context (e.g., Cycowicz et al., 2001), but not for core factual content.

Method

Participants

A total sample of 109 children and adolescents took part in the study (54 girls, 55 boys). They were recruited from different French cities and were all native French speakers. Most participants came from middle- to upper-class families and were overwhelmingly Caucasian. They were divided into seven age groups: sixteen 4- to 6-year-olds ($M = 4$ years 11 months, range = 4;3–5;3), fourteen 6-year-olds ($M = 6;8$, range = 6;3–6;11), fourteen 7-year-olds ($M = 7;4$, range = 7;0–7;8), fourteen 8-year-olds ($M = 8;6$, range = 8;0–8;11), eighteen 9-year-olds ($M = 9;6$, range = 9;0–9;11), fifteen 10- to 12-year-olds ($M = 10;8$, range = 10;1–11;4), and eighteen adolescents ($M = 15;1$, range = 14–15;10). All participants, except for those belonging to the youngest group, performed the Block Design subtest of the Wechsler
Intelligence Scale for Children (WISC–IV; Wechsler, 2004). The mean score of each group was within the norm, and there was no difference between the groups, $F(5, 81) < 1$. Ninety-two of the 109 participants were also assessed in terms of EF, and 77 in terms of episodic buffer capacity. These complementary tasks could not be administered to all the participants due to lack of time. Note, however, that none of the children had to be excluded on the grounds of poor comprehension or difficulty following the task instructions.

Exclusion criteria included neurological and psychiatric medical history, developmental learning disorders, and repetition of a year at school. The children’s participation was conditional upon approval by their head teachers and teachers, and their own willingness to take part in the experiment. Moreover, all parents provided written informed consent for their children’s participation in the study.

Procedure

Children were tested individually in a quiet room at their school on 2 different days within a 21-day window; each session lasted approximately 40 min. During the first session, participants performed the House Test, designed to assess EM abilities. During the second session, participants underwent an additional battery of cognitive tests probing nonverbal reasoning abilities (Block Design, WISC–IV; Wechsler, 2004), three elementary EFs that are seemingly crucial to children—updating, inhibition, and flexibility (Garon, Bryson, & Smith, 2008; see below)—and episodic buffer capacity. Tests were administered in a fixed sequence. Four- and 5-year-old children generally needed three sessions: one for the EM test and two for the other cognitive tests. Practice trials were administered before each task, to check that the instructions were fully understood.

Episodic Memory—The House Test

Overview. A board showing the front of a house was placed in front of the child. The experimenter began by explaining that he or she was going to describe a day in the life of a child (either a boy or a girl, depending on the participant’s gender) living in the house. The child was then described performing a variety of activities in the course of the day, each situated in time and space. There was then an interval of about 10 min, followed by a test phase in which the participants were asked to recall everything they could about the initial encoding phase in three ways: free recall, cued recall, and recognition. The procedure is described in greater detail below.

Materials. The board placed in front of the child measured $40 \times 62$ cm and depicted the front of a two-story house with nine areas (see Figure 1). Twenty-four $3.6 \times 3.6$ cm illustrations were also used: 12 evoked actions and 12 depicted objects. We selected actions that could be meaningfully linked to at least three different objects. Six additional pictures (three actions and three objects) were used for practice.

The day was divided into three periods (morning, afternoon, and evening), and participants were orally informed of this. For each period, six pictures were placed in front of the child (three images evoking actions, three objects), illustrating three different activities. After these pictures had been used to construct three sentences, they were taken away and replaced by six new illustrations for the next period. Immediately prior to the test, we conducted a trial run with six pictures to make sure the participant understood the instructions.

Encoding phase. The board was placed in front of the participant and six pictures were arranged in a fixed order below the board. The participant was told that he or she was going to play a game to check comprehension skills, but the subsequent retrieval tasks were not mentioned. Hence, encoding was incidental, allowing EM to be probed in a more naturalistic fashion. The experimenter described the different activities performed by the child living in the house in the course of a day. Nine activities were described in chronological order (i.e., from the moment the child got up in the morning to the moment he or she went to bed). Each activity consisted of three successive items of information: the temporal context (e.g., “after he got dressed”), the factual information (action associated with an object, e.g., “he poured water into,” “the aquarium”), and the spatial context (e.g., “in the bedroom”). Each time an activity was described by the experimenter, the participant had to pick up two of the six pictures depicting the factual information (e.g., pictures of a bottle of water and an aquarium) and put them in the right area. The temporal context had to be provided at the same time. This procedure encouraged the children to pay attention and allowed us to check that the different items of information had been correctly processed. Any errors were pointed out by the experimenter and the participant had to rectify them after being given the information again. If necessary, several
trials were performed until each item of information had been correctly processed. With the exception of a handful of participants, this procedure proved unnecessary and one repetition sufficed, regardless of the participants’ age.

**Test phase.** After approximately 10 min, the child was asked to recall as much of the information provided during the encoding phase as possible (free recall). To avoid any age-related differences in terms of knowledge about the experimenter’s expectations and to prevent information being omitted even though it had been memorized, participants were asked to systematically specify the spatial and temporal context of each factual item they recalled. When the participant could no longer recall any information, verbal cued recall was performed for any activity that had not been evoked during free recall. Cues consisted of the core action of the activity (in the previous example, “pouring water”). Each cue was given separately by the experimenter to help the participants remember each activity. Whenever possible, the latter had to recall both the object and the spatiotemporal context associated with this cue (“At a certain point, the child poured some water into something. Can you remember what he/she poured it into? In which area did he/she do it? When exactly did this take place?”). Finally, a forced-choice recognition test was administered for all the wrong or missing information. For each cue provided by the experimenter, the child was given the choice of three possible answers (in the previous example, factual recognition = “Did the child pour the water into a vase, an aquarium, or a plant pot?” spatial recognition = “Did this take place in the bedroom, the hall, or the study?” temporal recognition = “Did this happen after he/she had got dressed, after he/she had had a snack or after dinner?”). All the possible answers referred to information that had previously been encoded, to ensure that recall was not based merely on a feeling of familiarity. Lists of answers were constructed in such a way as to ensure that no item of information was used as a cue more than the others and that the correct item of information occupied each position in the list (first, second, third) the same number of times.

Three scores were calculated at the end of each test phase (free recall, cued recall, and recognition), according to the type of information evoked during that test phase, as well as the previous ones. For example, the cued recall scores were calculated by adding together the number of items of information retrieved during free and cued recall. In the same vein, the recognition scores were calculated by summing the items retrieved during free recall, cued recall, and recognition. The three scores were
(a) a factual score corresponding to the number of correctly remembered items of factual information, (b) a spatial context score corresponding to the number of factual items that were correctly remembered in association with the relevant spatial context, and (c) a temporal context score corresponding to the number of factual items that were correctly remembered in association with the relevant temporal context. The maximum score was always 9.

As the recall of factual and spatial information was based on visual cues, whereas temporal information was based on verbal information, an additional retrieval task was administered in order to probe temporal context memory using visual cues. In the temporal visual memory task, pictures of the nine previously encoded activities (action + object pictures) were given to the participants (see Figure 1). A timeline strip depicting the succession of different temporal contexts throughout the day was placed in front of them, too. The pictures illustrating the temporal contexts were separated by blank spaces where the participants could place the pictures according to their temporal context. The score reflected the number of pictures that were correctly situated in time, in relation to the three different periods (morning, afternoon, and evening). For example, 1 point was awarded if the picture of the first activity was placed in the first, second, or third gap. The maximum score was always 9 and the chance level was 3. Participants did not have to remember the specific temporal location (based on a verbal cue) in order to perform this task successfully, as they could also rely on visual cues, that is, by remembering when the activity took place in relation to the picture changes that occurred at the beginning of each period.

Short-Term Feature-Binding Assessment

Short-term feature binding was assessed with two computerized span tasks using material designed to be as comparable as possible to the EM task. Two tasks, adapted from Lorsbach and Reimer (2005) and Quinette et al. (2003), were used. In both tasks, the house picture featured in the EM test (with blanked out areas) served as the background. Participants had to memorize increasingly long strings of objects associated with a specific spatial context (area of the house) and recall them immediately afterward in the same order.

The unitized binding span task required participants to memorize readymade object–location associations, whereas the separated binding span task required them to memorize object–location associations that had not already been established (participants had to mentally bind the picture of an object displayed below the house with its location according to a color code; see Figure 2). The presentation of the stimuli was strictly controlled by SuperLab Pro 2.0: Objects were successively displayed on the screen for 2,000 ms, with an interstimulus interval of 500 ms. The end of the trial was symbolized by a question mark that remained on the screen for 500 ms. The participant then had to orally recall the sequence of encoded objects, pointing to their respective locations, in their order of appearance. Responses were deemed to be correct when all the information was correctly recalled. Three lists were provided for each sequence length, and testing stopped when participants failed on at least two of the three lists of a given length. The two short-term feature-binding scores (i.e., unitized and separated binding scores) corresponded to the longest list that was correctly recalled by the participant.

Executive Functions Assessment

Based on Miyake’s framework (Miyake et al., 2000), we focused our assessment on three specific functions: updating, inhibition, and shifting. These three EFs have been the focus of many developmental studies (for reviews, see Garon et al., 2008;
Updating. These abilities were assessed with a running span test (adapted from Morris & Jones, 1990). Word lists were played to the participants who, depending on the instructions, then had to recall either the last 2 or the last 3 words of the list, in the order in which they had heard them. There were two sets of 12 lists (recall of either last 2 or last 3 words), of different lengths (either 2, 4, 6, or 8 words, or 3, 5, 7, or 9 words). The participants were never told the length of the list beforehand, the 12 lists were randomly ordered, and this order was kept constant across participants. Nine bisyllabic words taken from the LEXIQUE French database were used (New, Pallier, Brysbaert, & Ferrand, 2004). The task was strictly controlled by SuperLab Pro 2.0: The pronunciation of each word lasted exactly 1,000 ms and the words were separated by a noiseless interval of 1,000 ms. Some of the lists did not require any updating and were included so that participants would realize that they could not afford to ignore any of the words. These particular lists were not taken into account for the updating score, which corresponded to the number of lists for which the participants correctly recalled the last two or three words (maximum = 18).

Cognitive inhibition. These abilities were assessed with a nonverbal Stroop task divided into three parts (Pennequin, Nanty, & Khomsi, 2004). In each one, participants had 60 s to process as many stimuli as possible. In the first part, they were shown pictures of animals all the same size and were asked to indicate how big or small the animals really were as fast as possible (i.e., butterflies are small animals, dogs are medium-sized ones, and elephants are large ones; see Figure 3). Each error was pointed out and had to be corrected before could move on to the next stimulus. In the second part, they were shown pictures of different-sized black squares and had to say whether they were small, medium, or large. The third part involved inhibition abilities. It required participants to give the actual sizes of animals which were deliberately shown in a conflicting size (e.g., butterfly pictures were either medium or large, but never small, as real-life butterflies are small). The number of stimuli processed in each part was recorded and an inhibition score was calculated: part 3 = [(part 1 * part 2) / (part 1 + part 2)].

Mental set shifting. It was assessed in an additional fourth part of the nonverbal Stroop task. The material was the same as in the third task, but the instructions differed: Participants had to alternate between perceptual processing (giving the size of the pictures) for one stimulus and semantic processing (giving the real-life size of the animals) for the other. The number of processed stimuli was recorded and a flexibility score was obtained by calculating the ratio of processed stimuli requiring flexibility to processed stimuli accompanied by simple instructions: part 4 * 100 / part 3.

All the scores were rated by two independent experts: the tester and a second rater. Interrater reliability, calculated for 20% of the scoring, showed agreement of more than 90%, whatever the measure.

Results

Effect of Age on the EM Test

The first set of analyses examined the developmental trajectory of each EM component (factual, spatial, temporal) in the three different test phases.

Figure 3. Examples of stimuli used in each of the trials of the four tasks (the first four correct responses are shown in italics; S = small; M = medium; L = large).
(free recall, cued recall, recognition). For each test phase, we conducted three successive analyses of variance (ANOVAs) on the different scores (as the contextual scores depended on the amount of factual information recalled, repeated measures of variance could not be used to examine the age group effect on all three scores simultaneously). First, we conducted a one-way ANOVA on the factual score, with age group as a between-participants factor. Second, we compared the effects of age on both types of context in a 7 (age group) × 2 (type of context) ANOVA with repeated measures on the second factor. Any significant effect revealed by the ANOVAs was followed by post hoc tests (Tukey’s honestly significant difference method) and the level of significance was set at $p < .05$ for all analyses. The effect size in each ANOVA was also reported, in order to indicate the strength of the relation independently of sample size (partial eta-squared values).

Figure 4 shows the mean scores of each age group, according to task (free recall, cued recall, recognition) and type of information (factual, factual + spatial context, factual + temporal context). The ANOVA performed on the free recall phase yielded a highly significant effect of age for factual information, in that performance improved with age, $F(6, 102) = 8.36, p < 10^{-5}, \eta_p^2 = .33$. Post hoc Tukey tests indicated that mean performance did not vary significantly either between the three youngest groups or between the three oldest groups, but that 4- to 6-year-old children recalled significantly less information than children aged 8 or more ($p < .001$ to $p < .0001$) and 7-year-old children recalled significantly less information than children aged 9 or more ($p < .05$ to $p < .001$). Context scores also increased with age, $F(6, 102) = 10.72, p < 10^{-6}, \eta_p^2 = .39$, and followed quite a similar pattern: four- to 6-year-old and 7-year-old children recalled the context less often than children aged 9 years or more ($p < .0001$). Finally, the spatial context was recalled better than the temporal context, $F(1, 102) = 71.81, p < 10^{-5}, \eta_p^2 = .41$, and this pattern did not vary with age, no interaction between age group and context, $F(6, 102) = 1.33, ns$.

ANOVA conducted on the cued recall scores revealed an effect of age group for the factual information, $F(6, 102) = 5.48, p < 10^{-4}, \eta_p^2 = .25$. The results of the post hoc tests showed that performance increased steadily, with 4- to 6-year-old children recalling less information than children aged 9–12 years ($p < .001$ to $p < .0001$) and 6-year-olds recalling less information than 10- to 12-year-olds. No other post hoc test reached significance. Like free recall, cued recall was characterized by an effect of age on context recall, $F(6, 102) = 10.33, p < 10^{-6}, \eta_p^2 = .38$. Spatial context was recalled better than temporal context, $F(1, 102) = 183.11, p < 10^{-6}, \eta_p^2 = .64$, and there was no interaction between age group and context, $F(6, 102) = 1.61, ns$.

For the recognition task, the analysis revealed a main effect of group, $F(6, 101) = 6.18, p < 10^{-4}, \eta_p^2 = .27$, reflecting the fact that the youngest group recognized fewer items than participants significantly less information than children aged 8 or more ($p < .001$ to $p < .0001$) and 7-year-old children recalled significantly less information than children aged 9 or more ($p < .05$ to $p < .001$). Context scores also increased with age, $F(6, 102) = 10.72, p < 10^{-6}, \eta_p^2 = .39$, and followed quite a similar pattern: four- to 6-year-old and 7-year-old children recalled the context less often than children aged 9 years or more ($p < .0001$). Finally, the spatial context was recalled better than the temporal context, $F(1, 102) = 71.81, p < 10^{-5}, \eta_p^2 = .41$, and this pattern did not vary with age, no interaction between age group and context, $F(6, 102) = 1.33, ns$.

ANOVA conducted on the cued recall scores revealed an effect of age group for the factual information, $F(6, 102) = 5.48, p < 10^{-4}, \eta_p^2 = .25$. The results of the post hoc tests showed that performance increased steadily, with 4- to 6-year-old children recalling less information than children aged 9–12 years ($p < .001$ to $p < .0001$) and 6-year-olds recalling less information than 10- to 12-year-olds. No other post hoc test reached significance. Like free recall, cued recall was characterized by an effect of age on context recall, $F(6, 102) = 10.33, p < 10^{-6}, \eta_p^2 = .38$. Spatial context was recalled better than temporal context, $F(1, 102) = 183.11, p < 10^{-6}, \eta_p^2 = .64$, and there was no interaction between age group and context, $F(6, 102) = 1.61, ns$.

For the recognition task, the analysis revealed a main effect of group, $F(6, 101) = 6.18, p < 10^{-4}, \eta_p^2 = .27$, reflecting the fact that the youngest group recognized fewer items than participants

![Figure 4](chart.png)

Figure 4. Chart showing mean recall scores on the episodic memory task, according to age group, task (free recall, cued recall, recognition), and type of information (factual information, spatial context associated with factual information, temporal context associated with factual information, complex = factual information associated with both contexts).

Note. The dotted line represents the maximum score.
aged 7 or more \((p < .001\) to \(p < .0001\)). Moreover, the performances of the 7- to 16-year-old participants were characterized by a ceiling effect, as they did not differ significantly from the maximum score (one-sample \(t\)-test), \(t_{8\text{ years old}}(13) = 1.96, ns\); \(t_{8\text{ years old}}(13) = 1.95, ns\); \(t_{9\text{ years old}}(17) = 1.89, ns\); \(t_{10\text{ to 12 years old}}(14) = 1.44, ns\); \(t_{14\text{ to 16 years old}}(17) = 1.46, ns\). Recognition of the context also improved with age, \(F(6, 102) = 15.09, p < 10^{-5}, \eta_p^2 = .47\), with 4- to 6-year-old children recalling less context than all the other groups except the 7-year-olds who, in turn, recalled less information than children aged 9 or more \((p < .0001)\). Again, the effect of context was significant, with the spatial context being recognized more often than the temporal one, \(F(1, 102) = 53.81, p < 10^{-5}, \eta_p^2 = .34\). However, in contrast to the other two phases, the interaction between age group and context was significant in the recognition task, \(F(6, 102) = 2.31, p < .05, \eta_p^2 = .12\). Post hoc tests revealed that recognition of the spatial context was less difficult than recognition of the temporal context for young children (4- to 6-year-old and 7-year-old children, \(p < .0001\)), whereas there was no statistical difference between spatial and temporal context from 8 years onward.

The ANOVA conducted on the temporal visual memory task revealed an effect of age group, \(F(6, 86) = 9.45, p < 10^{-5}, \eta_p^2 = .38\) (see Figure 5). The results of the post hoc tests showed that less information was correctly situated in time by 4- to 6-year-old children than by the other groups \((p < .06\) to \(p < .001)\). Moreover, children aged 6–8 years differed significantly from the 10- to 12-year-old group. No other post hoc test reached significance. Note that the 4- to 6-year-old participants accurately situated only three actions on average, which was not significantly different from chance level (one-sample \(t\)-test), \(t(13) = 0, ns\). Conversely, performances were significantly above chance levels from 6 years upward \((p < .001\) to \(p < 10^{-3}\)).

In summary, these results revealed a large age-related increase in EM abilities between 4 years of age and adolescence, with different developmental trajectories according to the particular EM component being recalled and the type of task being performed. In short, the ability to recall factual information in the free and cued recall tasks increased gradually with age, whereas the recognition task allowed all participants up to 7 years to retrieve the same amount of factual information. By contrast, the spatial context and, a fortiori the temporal context, appeared to rely on complex processes that remained age sensitive, whatever the task and the number of self-initiated retrieval processes involved.

**Mechanisms Underlying the Development of the Different EM Components**

The second set of analyses sought to identify the mechanisms underlying the age-related variance in the EM free recall scores. The EFs and the short-term feature-binding processes were closely intercorrelated (see Table 1), and correlated with the EM scores (with \(r_s\) ranging from .16 to .47, \(p < .05\) to \(p < 10^{-5}\)). In order to test the potential mediating effect of EFs and short-term feature binding on the three EM scores, six hierarchical regressions were conducted: two for the factual score, two for the spatial score, and two for the temporal score. First, we successively forced the two factors (EFs and short-term feature binding) to examine whether they accounted for a significant proportion of variance in EM scores. Then, we entered age group as a predictor, and examined if its contribution was reduced to a nonsignificant level. The reasoning was that if age was no longer a significant predictor of EM score after controlling for cognitive abilities, then age-related differences in EM must be due to differences in these abilities, which are known to improve with age (for a comparable statistical method, see Bugaiska et al., 2007; Picard, Rebuffe-Veilleux, Eustache, & Piolino, 2009). In the first analysis, we examined whether the EFs continued to account for a significant amount of EM scores’ variance when short-term feature-binding measures were entered first. In the second analysis, we examined the opposite relation: Do short-term feature-binding abilities account for a significant propor-

Figure 5. Graph showing mean scores on the temporal visual memory task, according to age group.

Note. The dotted line represents the maximum score and the solid line the chance level.
tion of EM scores’ variance, when EFs measures are entered prior? As recommended by Bryan and Luszcz (1996), the age group variable rather than individual age was used in these statistical analyses.

The results are shown in Table 2. First, the results showed that short-term feature binding was a significant predictor of the factual score (11% of the variance), whether it was entered before (Analysis 1) or after (Analysis 2) EFs. By contrast, EFs did not contribute significantly to the factual variance (Analyses 1 and 2). Analyses 1 and 2 showed that the EFs and short-term feature-binding abilities contributed in a similar proportion to the spatial score (between 7% and 10%, depending on the analysis). But EFs contributed more to the temporal score than short-term binding abilities: EFs significantly predicted a large part of variance, whether entered first (19% in Analysis 1), or after short-term binding abilities (27% in Analysis 2). In contrast, short-term feature-binding abilities accounted for 16% of the variance when entered first (Analysis 1), and for only 8% once EFs had been controlled for. Second, the analyses showed that after controlling for the two factors (short-term binding abilities and EFs), age added only 5% to the factual score variance and 4% to the spatial context variance, which

### Table 2

**Hierarchical Multiple Regression Analyses Predicting Episodic Memory Scores Obtained During Free Recall Using Short-Term Feature Binding, Executive Functions and Age as Predictors**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Factual</th>
<th></th>
<th>Spatial</th>
<th></th>
<th>Temporal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AR²</td>
<td>β</td>
<td>AR²</td>
<td>β</td>
<td>AR²</td>
</tr>
<tr>
<td><strong>Analysis 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term feature binding</td>
<td>.11*</td>
<td>.09*</td>
<td></td>
<td></td>
<td>.16**</td>
</tr>
<tr>
<td>Unitized binding</td>
<td>.34*</td>
<td></td>
<td>.36*</td>
<td></td>
<td>.51**</td>
</tr>
<tr>
<td>Separated binding</td>
<td>-.01</td>
<td></td>
<td>-.10</td>
<td></td>
<td>-.22</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive functions</td>
<td>.04</td>
<td>.08*</td>
<td>.19**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inhibition</td>
<td>-.08</td>
<td></td>
<td>-.02</td>
<td></td>
<td>.25</td>
</tr>
<tr>
<td>Shifting</td>
<td>.16</td>
<td></td>
<td>.28*</td>
<td></td>
<td>.44*</td>
</tr>
<tr>
<td>Updating</td>
<td>.08</td>
<td></td>
<td>-.07</td>
<td></td>
<td>.11</td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.05</td>
<td>.28</td>
<td>.25</td>
<td></td>
<td>.07*</td>
</tr>
<tr>
<td>Total R²</td>
<td>.20</td>
<td>.21</td>
<td>.25</td>
<td></td>
<td>.07*</td>
</tr>
<tr>
<td><strong>Analysis 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive functions</td>
<td>.04</td>
<td>.10*</td>
<td>.27**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inhibition</td>
<td>.11</td>
<td></td>
<td>.04</td>
<td></td>
<td>.28*</td>
</tr>
<tr>
<td>Shifting</td>
<td>.19</td>
<td></td>
<td>.31*</td>
<td></td>
<td>.47***</td>
</tr>
<tr>
<td>Updating</td>
<td>.06</td>
<td></td>
<td>.15</td>
<td></td>
<td>.15</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term feature binding</td>
<td>.11*</td>
<td>.07*</td>
<td>.08*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unitized binding</td>
<td>.36**</td>
<td></td>
<td>.35*</td>
<td></td>
<td>.40*</td>
</tr>
<tr>
<td>Separated binding</td>
<td>.10</td>
<td></td>
<td>-.02</td>
<td></td>
<td>-.29</td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.05</td>
<td>.28</td>
<td>.25</td>
<td></td>
<td>.07*</td>
</tr>
<tr>
<td>Total R²</td>
<td>.20</td>
<td>.21</td>
<td>.25</td>
<td></td>
<td>.07*</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
was not significant. In contrast, age continued to predict temporal context performance adding 7% to the variance after EFs and short-term feature-binding abilities had been controlled for.

In summary, short-term binding abilities appeared to be a reliable predictor of variance in all EM scores, whereas EF only contributed to contextual scores. Further, age added a significant contribution to EM variance, over and above EFs and short-term binding abilities only for temporal context.

Discussion

The purpose of this study was to further characterize the development of EM using an ecological task designed specifically for this purpose. Based on a rich and updated definition of EM (Tulving, 2001), we compared the developmental trajectories of its three main components (core factual content, and spatial and temporal contexts) across a wide age range, from early childhood to adolescence, and investigated the underlying sources of these age-related changes. On the whole, our data confirm the multidimensional nature of the EM system and its slow pace of maturation, dependent upon the development of various separable components. They argue in favor of the notion that the development of each EM component across childhood is contingent upon short-term feature-binding abilities, with context memory also relying on EFs. That said, although the ability to recall temporal context was associated with these two cognitive factors, age contributed significantly above these two cognitive factors.

How Do the Different EM Components Develop?

As concerns the core factual information, our results demonstrated that children could correctly remember a great deal of information from the age of 4 years onward, even though its encoding was incidental. More specifically, we found that the ability to remember core factual information about the activities improved nonlinearly across childhood, and that three main periods of change could be identified: This ability increased markedly during preschool years (see also Newcombe et al., 2007), continued to improve slightly between the ages of 6 and 9 years (see also Waber et al., 2007), and reached maturity at around 9 years. Each of the three above periods of change was influenced by a different mechanism. The poorer preschool performance reflected poorer encoding abilities. Even when the retrieval process were greatly facilitated (recognition test), the 4- to 6-year-old children remained unable to access all the factual information, contrasting with the children over the age of 7. By contrast, improvements in retrieval seemed mainly to underlie the increase in abilities between 6 and 9 years, as children of this age differed only slightly from older children on the free and cued recall tasks, and performed similarly on the recognition task. Finally, no major change was observed from 9 years onward: All the children encoded all the factual information (all the items were correctly recognized) and spontaneously retrieved half of it, suggesting that the processes of encoding and retrieving core factual content reach maturity at around 9 years. In sum, we showed that memory for factual information increases nonlinearly between 4 and 16 years, subtended by successive improvements in encoding and retrieval abilities.

The younger children found it extremely difficult to remember the encoding context of the memorized activities, in line with previous findings (Gulya et al., 2002; Newcombe et al., 2007). Furthermore, we were able to show for the first time within the same sample that the two types of context recall are differently sensitive to age, arguing in favor of two independent processes (see Russell & Thompson, 2003, for a similar finding in infancy; see also Ekstrom & Bookheimer, 2007). More specifically, we found that both contexts followed comparable developmental trajectories in free and cued recall, with the spatial context being recognized more often than the temporal one. In recognition, the 4- to 7-year-old children still experienced greater difficulty accessing the temporal context than the spatial one. By contrast, children over 8 extensively benefited from the recognition phase for the temporal context, leading to similar performance for the two contexts. Memorizing the temporal context of events therefore appears to be a complex process, especially for young children (for comparable results viewed from an autobiographical memory perspective, see Piolino et al., 2007). Their deficits seem to stem from inefficient encoding processes, whereas difficulty accessing the spatial context of events in childhood seems to arise mainly from retrieval deficits.

Nonetheless, it is unclear whether the specific issue we observed with temporal material was due to the material we used, or whether reflected a genuine developmental issue. The encoding of factual and spatial information was conveyed by visual cues, whereas the temporal information was only communicated verbally. The specific issue with
temporal material may thus have been due to the material. However, the results of the temporal visual memory task do not support this hypothesis. Successfully performing this task did not require recall of the specific temporal location (based on a verbal cue), as participants could rely on visual cues (during the encoding phase, the experimenter changed the pictures at the beginning of each period). Regardless of task, our results therefore point to a specific developmental issue concerning the memorization of the temporal context associated with core factual information.

How Do Short-Term Feature Binding and EF Contribute to EM Development?

While it is certainly worthwhile identifying the mechanisms that underpin EM development, it is even more worthwhile elucidating the reasons for the changes. In this respect, we found that the development of the different EM components relies on different processes. Consistent with our hypothesis, our results substantiate the view that short- and long-term feature-binding abilities are closely linked, seemingly due to the fact that both involve remembering associations between different items of information. Thus, the gradual increase in short-term feature-binding abilities across childhood (see also Cowan et al., 2006; Lorsbach & Reimer, 2005) appears to contribute to EM recall. Remarkably, we observed this link whatever the EM component. According to Baddeley’s working memory (WM) framework (2000), short-term abilities can be interpreted as reflecting the work of the episodic buffer. WM is regarded as a short-term memory system comprising multiple components, including the episodic buffer, which is involved in short-term feature-binding processes. As the episodic buffer allows multidimensional representations to be integrated and stored for short periods, it is assumed to play an important role in encoding information in long-term EM (see also Eustache & Desgranges, 2008; Reinitz & Hannigan, 2004) and later in retrieving information from it (Baddeley, 2000). Unfortunately, our results did not allow us to disentangle these two effects and further studies are therefore required. Given that the material used in the house test and in the short-term feature-binding tasks was rather similar, this link could be regarded as spurious and simply reflecting the similarity of the material we used. However, it was not the fact as only one of the short-term feature-binding tasks was associated with the EM scores. For core factual information, short-term binding abilities were even the main predictor, mediating all of the previously observed effect of age. Further, the effect of age on the two contextual EM components was only partly mediated by short-term feature-binding abilities.

Indeed, EFs were also involved in contextual recall, in accordance with our hypothesis. Like Cycoiwicz et al. (2001), our results indicate that EFs are crucial to the development of EM abilities, allowing the encoding of rich EMs and contributing to their subsequent spontaneous recall. Previous studies have emphasized the role of EFs in the sense of time, which is crucially involved in EM (e.g., Mantylä, Carelli, & Forman, 2007). The continuous improvement in EFs in our sample supports this hypothesis and extends it to the spatial context. These results also build on the recent findings of Picard et al. (2009), who adopted an autobiographical perspective. These authors found that the effect of age on the ability to retrieve personal memories was mainly linked to the effect of age on EFs.

Thus, the development of short-term feature-binding abilities and EFs appears to be crucial for factual and spatial recall in childhood, accounting for almost all of the development of these abilities. In contrast, the effect of age on recall of temporal context that we observed was only partially explained by short-term binding abilities and EFs. Other age-sensitive cognitive factors may therefore also be involved, such as the development of strategies (Schneider & Pressley, 1997), metamemory (Bjorklund & Douglas, 1997; DeMarie & Ferron, 2003) and semantic knowledge (Robertson & Köhler, 2007). In our task, the temporal context referred to times of the day and required conceptual knowledge about conventional time patterns (i.e., location-based temporal process) that, according to Friedman (2005), may undergo great developmental change. Thus, temporal context may represent particular difficulties insofar as it refers to abstract concepts that are difficult to represent using perceptual representations and visual mental imagery, unlike spatial context and factual information.

In conclusion, the novel finding of the present study is its demonstration of differential developmental trajectories of the core and contextual components of EM in childhood and adolescence, subtended by different mechanisms. Even if our results are in line with our predictions and with the literature, the interpretations must nonetheless be viewed with caution, as they are only based on single tasks. Our findings open up new avenues for research, as further studies are needed to replicate and extend our findings using different tasks, and
to unravel the complete set of mechanisms responsible for this development.

References


